

Search completed: June 13, 2001, 14:14:30
Job time: 371 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:10:36 ; Search time 130.61 seconds
(without alignments)
4.209 Million cell updates/sec

Title: PCT-US01-05825A-2

Perfect score: 45

Sequence: 1 GRVCVQDG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_67:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	42	93.3	399	2 B82197	zona occludens tox
2	42	93.3	399	2 A43864	zonula occludens t
3	36	80.0	516	2 B81684	exodeoxyribonuclea
4	36	80.0	516	2 F71529	probable exodoxyl
5	35	77.8	322	2 S41400	aspartic proteinase
6	35	77.8	475	2 T01352	hypothetical prote
7	34	75.6	796	2 T23238	hypothetical prote
8	33	73.3	145	2 S74292	hypothetical prote
9	33	73.3	185	2 B82065	conserved hypothet
10	33	73.3	196	2 F71525	hypothetical prote
11	33	73.3	196	2 H81681	maf protein TC0628
12	33	73.3	340	2 S69194	N4-(beta-N-acetyl
13	33	73.3	440	2 T43198	nitrilase/Phit pro
14	33	73.3	471	2 G82083	glutamate synthase
15	33	73.3	472	2 G65112	glutamate synthase
16	33	73.3	477	2 G83017	glutamate synthase
17	33	73.3	494	2 H82523	glutamate synthase
18	33	73.3	508	2 T54537	anthranilate synth
19	33	73.3	524	2 JN0506	ATP-stimulated glu
20	33	73.3	582	2 T28912	probable homeobox
21	33	73.3	644	2 G65072	hypothetical prote
22	33	73.3	659	2 C65022	yifg protein - Esc
23	33	73.3	676	2 A45984	sperm-binding glyc
24	33	73.3	745	1 A48833	sperm-binding glyc
25	33	73.3	825	1 GLHQ	beta-glucosidase (
26	33	73.3	979	1 JC2349	protein-tyrosine-p
27	33	73.3	996	2 I48721	PTP 35 protein - m
28	33	73.3	1012	2 T41940	DNA polymerase - h
29	32	71.1	224	2 C72010	conserved hypothet

ALIGNMENTS

RESULT 1

B82197

zona occludens toxin VC1458 [Imported] - Vibrio cholerae (group O1 strain N16961)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000

C:Accession: B82197

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R. Chardson, D.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: B82197

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-399 <HEI>

A:Cross-references: GB:AE004224; GB:AE003852; MID:g9655952; PIDN:AAF94615.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1458

A:Map position: 1

Query Match 93.3%; Score 42; DB 2; Length 399;

Best Local Similarity 87.5%; Pred. No. 1.1;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRVCVQDG 8

Db 291 GRVCVQDG 298

RESULT 2

A43864

zonula occludens toxin - Vibrio cholerae

C:Species: Vibrio cholerae

C:Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999

C:Accession: A43864

R:Baudry, B.; Fasano, A.; Kettley, J.; Kaper, J.B.

Infect. Immun. 60, 428-434, 1992

A:Title: Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.

A:Reference number: A43864; MUID:92112300

A:Accession: A43864

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-399 <BAU>

A:Cross-references: GB:M83563; MID:g155314; PIDN:AAA27582.1; PID:g155315

A>Note: sequence extracted from NCBI backbone (NCBIN:77488, NCBIp:77491)

Query Match 93.3%; Score 42; DB 2; Length 399;

Best Local Similarity 87.5%; Pred. No. 1.1;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVQDG 8
||:|||||
DB 291 GRCLVQDG 298

RESULT 3
B81684
exodeoxyribonuclease, large chain TC0605 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000
A:Accession: B81684
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255
A:Accession: B81684
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-516 <TEF>
A:Cross-references: GB:AE002329; GB:AE002160; NID:g7190640; PIDN:AAF39436.1; PID:g719064
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0605
C:Superfamily: exodeoxyribonuclease VII

Query Match 80.0%; Score 36; DB 2; Length 516;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVCVQDG 8
||:|||||
DB 493 RVCLQDG 499

RESULT 4
F71529
probable exodeoxyribonuclease VII - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 18-Aug-2000
A:Accession: F71529
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A:Reference number: A71570; MUID:99000809
A:Accession: F71529
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-516 <ARN>
A:Cross-references: GB:AE001305; GB:AE001273; NID:g3328737; PIDN:AAC67922.1; PID:g332874
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: xseA
C:Superfamily: exodeoxyribonuclease VII

Query Match 80.0%; Score 36; DB 2; Length 516;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVCVQDG 8
||:|||||
DB 493 RVCLQDG 499

RESULT 5
S41400
aspartic proteinase (EC 3.4.23.-) - wild cabbage (fragment)
N:Alternate names: aspartyl protease
C:Species: Brassica oleracea (wild cabbage)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 29-Aug-1997
C:Accession: S41400
R:Fujikura, Y.; Karssen, C.M.
submitted to the EMBL Data Library, January 1994
A:Description: Cauliflower cDNA with sequence homology to gastric proteases.
A:Reference number: S41400
A:Accession: S41400
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-322 <FUJ>
A:Cross-references: EMBL:X77260
C:Comment: The pair of saposin repeat homology domains tagged SAP1 and SAP2 represent
C:Superfamily: oryzasin; saposin repeat homology
C:Keywords: aspartic proteinase; hydrolase
F:99-144/Domain: saposin repeat homology #status atypical <SAP1>
F:153-203/Domain: saposin repeat homology #status atypical <SAP2>
F:73/Active site: Asp #status predicted

Query Match 77.8%; Score 35; DB 2; Length 322;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVQDG 7
||:|||||
DB 316 GRVCVKD 322

RESULT 6
T01352
hypothetical protein F6N15.2 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 14-May-1999
C:Accession: T01352
R:Ryan, E.; Edwards, J.; Pape, K.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of A. thaliana F6N15.
A:Reference number: Z14297
A:Accession: T01352
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-475 <RYA>
A:Cross-references: EMBL:AF069299; NID:g3193311; PID:g3193319
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Introns: 194/1
A:Note: F6N15.2

Query Match 77.8%; Score 35; DB 2; Length 475;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVCVQDG 8
||:|||||
DB 208 RICVEDG 214

RESULT 7
T23238
hypothetical protein T05H10.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T23238; T24560
R:Lightning, J.
submitted to the EMBL Data Library, January 1995
A:Reference number: Z19713
A:Accession: T23238
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-796 <WIL>
A:Cross-references: EMBL:Z47811; PIDN:CAA87788.1; GSPDB:GN00020; CESP:T05H10.7

A:Experimental source: clone K02C4
R:Thomas, K.
submitted to the EMBL Data Library, January 1995
A:Reference number: Z19907
A:Accession: T24560
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-796 <W12>
A:Cross-references: EMBL:Z47812; PIDN:CAA87796.1; GSPDB:GNO0020; CESP:T05H10.7
C:Genetics:
A:Gene: CESP:T05H10.7
A:Map position: 2
A:Introns: 77/1; 128/2; 163/1; 249/3; 559/3; 561/3; 648/3; 776/3

Query Match 75.6%; Score 34; DB 2; Length 796;
Best Local Similarity 75.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVCVQDG 8
| | | | |
Db 198 GRKCVSDG 205

RESULT 8
S74292
hypothetical protein YCR068w-a - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 22-Oct-1999
C:Accession: S74292
R:Wedler, H.; Wambutt, R.
submitted to the Protein Sequence Database, September 1996
A:Reference number: S74288
A:Accession: S74292
A:Molecule type: DNA
A:Residues: 1-145 <WED>
A:Cross-references: EMBL:X59720; GSPDB:GNO0003; MIPS:YCR068w-a
C:Genetics:
A:Gene: MIPS:YCR068w-a
A:Map position: 3R
C:Superfamily: Saccharomyces hypothetical protein YCR068w-a

Query Match 73.3%; Score 33; DB 2; Length 145;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GRVCVQD 7
| | | | |
Db 40 GRVCVYD 46

RESULT 9
B82065
conserved hypothetical protein VC2524 [imported] - Vibrio cholerae (group O1 strain N169)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: B82065
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, R.
L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: B82065
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-185 <HEI>
A:Cross-references: GB:AE004322; GB:AE003852; NID:g9657104; PIDN:AAF95666.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2524

A:Map position: 1
C:Superfamily: hypothetical protein H11679

Query Match 73.3%; Score 33; DB 2; Length 185;
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RVCVQDG 8
| | | | |
Db 136 RVCVADG 142

RESULT 10
F71525
hypothetical protein CT349 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C:Accession: F71525
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
A:Reference number: A71570; MUID:99000809
A:Accession: F71525
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-196 <ARN>
A:Cross-references: GB:AE001308; GB:AE001273; NID:g3328766; PIDN:AAC67944.1; PID:g332
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: CT349
C:Superfamily: septum formation protein maf

Query Match 73.3%; Score 33; DB 2; Length 196;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVCVQDG 8
| | | | |
Db 150 GGYCVQDG 157

RESULT 11
H81681
maf protein TC0628 [imported] - Chlamydia muridarum (strain N199)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C:Accession: H81681
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: H81681
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-196 <TEF>
A:Cross-references: GB:AE002331; GB:AE002160; NID:g7190662; PIDN:AAF39457.1; PID:g719
A:Experimental source: strain N199 (MoPn)
C:Genetics:
A:Gene: TC0628
C:Superfamily: septum formation protein maf

Query Match 73.3%; Score 33; DB 2; Length 196;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVCVQDG 8
| | | | |
Db 150 GGYCVQDG 157

RESULT 12

S69194
 N4-(beta-N-acetylglucosaminyl)-L-asparaginase (EC 3.5.1.26) precursor [validated] - Flav
 C:Species: Flavobacterium meningosepticum
 C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 26-May-2000
 C:Accession: S69194
 R:Tarentino, A.L.; Quinones, G.; Hauer, C.R.; Changchien, L.M.; Plummer Jr., T.H.
 Arch. Biochem. Biophys. 316, 399-406, 1995
 A:Title: Molecular cloning and sequence analysis of Flavobacterium meningosepticum glyco
 A:Reference number: S69194; MUID:95142655
 A:Accession: S69194
 A:Molecule type: DNA
 A:Residues: 1-340 <TAR>
 A:Cross-references: EMBL:U08028; NID:g555667; PIDN:AAA68868.1; PID:g555668
 C:Complex: heterodimer; alpha (PIR:S69194) and beta (PIR:S69194) chain [validated; MUID:
 C:Function:
 A:Description: EC 3.5.1.26 [validated; MUID:95142655]
 C:Superfamily: N4-(beta-N-acetylglucosaminyl)-L-asparaginase
 C:Keywords: heterodimer; signal sequence #status predicted <SIG>
 F:1-45/Domain: signal sequence #status predicted <SIG>
 F:46-196/Product: N4-(beta-N-acetylglucosaminyl)-L-asparaginase alpha chain #status pred
 F:197-340/Product: N4-(beta-N-acetylglucosaminyl)-L-asparaginase beta chain #status expe
 F:197/Active site: Thr #status predicted

Query Match

73.3%; Score 33; DB 2; Length 340;

Best Local Similarity 62.5%; Pred. No. 54;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVDG 8

| | | | |

Db 312 GAYCIODG 319

RESULT 13

T43198
 nitrilase/Fh1t protein - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T43198
 R:Pekarsky, Y.; Campiglio, M.; Siprashvili, Z.; Druck, T.; Sedkov, Y.; Tillib, S.; Draga
 proc. Natl. Acad. Sci. U.S.A. 95, 8744-8749, 1998
 A:Title: Nitrilase and Fh1t homologs are encoded as fusion proteins in Drosophila melanc
 A:Reference number: 223332; MUID:98337986
 A:Accession: T43198
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-440 <PEK>
 A:Cross-references: EMBL:AF069986; NID:g3228663; PIDN:AAC39136.1; PID:g3228664
 C:Genetics:
 A:Gene: NitFh1t

Query Match

73.3%; Score 33; DB 2; Length 440;

Best Local Similarity 83.3%; Pred. No. 69;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 VCVDQD 8

: | | | | |

Db 376 ICVDQD 381

RESULT 14

G82083
 glutamate synthase, small subunit VC2377 [imported] - Vibrio cholerae (group O1 strain N
 C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000
 C:Accession: G82083
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833

A:Accession: G82083

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-471 <HEI>

A:Cross-references: GB:AE004308; GB:AE003852; NID:g9656946; PIDN:AAF95520.1; GSPDB:GN
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2377

A:Map position: 1

C:Superfamily: glutamate synthase small chain

Query Match

73.3%; Score 33; DB 2; Length 471;

Best Local Similarity 85.7%; Pred. No. 73;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQD 7

| | | | |

Db 96 GRVCPQD 102

RESULT 15

G65112

glutamate synthase (NADPH) (EC 1.4.1.13) small chain - Escherichia coli

C:Species: Escherichia coli

C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 29-Sep-1999

C:Accession: G65112; B29617

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
 .A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: G65112

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-472 <BLAT>

A:Cross-references: GB:AE000400; GB:U00096; NID:g2367203; PIDN:AAC76245.1; PID:g17896
 A:Experimental source: strain K-12, substrain MG1655

R:Oliver, G.; Gosset, G.; Sanchez-Pescador, R.; Lozoya, E.; Ku, L.M.; Flores, N.; Bec
 Gene 60, 1-11, 1987

A:Title: Determination of the nucleotide sequence for the glutamate synthase structur

A:Reference number: A91585; MUID:88152492

A:Contents: K12

A:Accession: B29617

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-37, 'ARP', 41, 'RLTAACRA', 52-122, 'K', 124-173, 'C', 175-256, 'CTQRCRSSSTPNS
 A:Cross-references: GB:M18747; NID:g146207; PIDN:AAA3905.1; PID:g146209

C:Genetics:

A:Gene: gltD

A:Map position: 69 min

C:Superfamily: glutamate synthase small chain

C:Keywords: glutamate biosynthesis; NADP; oxidoeductase

Query Match

73.3%; Score 33; DB 2; Length 472;

Best Local Similarity 85.7%; Pred. No. 73;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVQD 7

| | | | |

Db 95 GRVCPQD 101

Search completed: June 13, 2001, 14:10:37

Job time: 138 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:21:39 ; Search time 74.44 Seconds
(without alignments)
3.681 Million cell updates/sec

Title: PCT-US01-05825A-2
Sequence: 1 GRVCVQDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	93.3	399	1	ZOT_VIBCH
2	37	83.2	1120	1	P38442 vibrio chol
3	34	75.6	70	1	Q85428 rat cytochrome
4	34	75.6	785	1	NUSG CGDAB
5	34	75.6	796	1	SM3E_CHICK
6	34	75.6	1171	1	YR57_CAEEL
7	34	75.6	1171	1	DPOL_HSV1
8	33	73.3	340	1	ASPG_FLAME
9	33	73.3	471	1	GLFD_ECOLI
10	33	73.3	508	1	TREPE_SYNY3
11	33	73.3	524	1	GLPK_MOUSE
12	33	73.3	524	1	GLPK_MOUSE
13	33	73.3	639	1	YGFY_ECOLI
14	33	73.3	659	1	AECA_ECOLI
15	33	73.3	666	1	ZP2_RABIT
16	33	73.3	745	1	ZP2_HUMAN
17	33	73.3	745	1	ZP2_HUMAN
18	33	73.3	825	1	BGLS_HANAN
19	33	73.3	979	1	PRPN_BOVIN
20	33	73.3	979	1	PRPN_HUMAN
21	33	73.3	979	1	PRPN_HUMAN
22	33	73.3	983	1	PRPN_MOUSE
23	33	73.3	1012	1	DPOL_HSV1
24	32	71.1	480	1	GLTD_AZOB
25	32	71.1	1569	1	GLT3_XENLA
26	31	68.9	122	1	YFPO_ECOLI
27	31	68.9	258	1	HTPX_METH
28	31	68.9	316	1	HP23_VZVD
29	31	68.9	342	1	CHSP_SCHTA
30	31	68.9	342	1	HOPK_AZOVI
31	31	68.9	348	1	HOPK_AZOVI
32	31	68.9	446	1	CRB3_SCHPO
33	31	68.9	496	1	PACR_MOUSE

34	31	68.9	511	1	C4A2_DROME	O9YMS8 drosophila
35	31	68.9	523	1	PACR_RAT	P32215 ratius norv
36	31	68.9	661	1	BAIH_EUBSP	P32370 eubacterium
37	31	68.9	715	1	ZP2_CANFA	P47983 canis fam11
38	31	68.9	716	1	ZP2_FELCA	P47984 felis silve
39	31	68.9	722	1	Y022_TREPA	O81066 treponema p
40	31	68.9	775	1	SM3E_HUMAN	O15041 homo sapien
41	31	68.9	775	1	SM3E_MOUSE	P70275 mus musculu
42	31	68.9	1035	1	DPOL_RHOM6	O71121 rhesus cyto
43	31	68.9	1242	1	DPOL_HCMVA	P08546 human cytom
44	31	68.9	1518	1	POLG_PPVEA	O01681 P genome po
45	31	68.9	3023	1	POLG_TYMW	P09814 t genome po

ALIGNMENTS

RESULT 1	ID	STANDARD	PRT	399 AA.
ZOT_VIBCH	1			
AC	P38442; Q9L706; Q9R3V6;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-2000 (Rel. 40, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	ZONA OCCLUDENS TOXIN (ZONULAR OCCLUDENS TOXIN).			
GN	ZOT OR VC1458.			
OS	Vibrio Cholerae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.			
OX	NCBI_Taxid=666;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CLASSICAL INABA 569B;			
RX	MEDLINE-92112300; PubMed-1730472;			
RA	Baudry B., Fasano A., Kettle J., Kaper J.B.;			
RT	"Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.";			
RL	Infect. Immun. 60:428-434(1992).			
RN	(2)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-KNH002;			
RA	Shin H.J., Park Y.C., Kim Y.C.;			
RT	"Cloning and nucleotide sequence analysis of the virulence gene cassette from Vibrio cholerae KNH002 isolated in Korea.";			
RL	Misatunurag Hojl 35:205-210(1999).			
RN	(3)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-0139-TOR OGAMA;			
RA	Zhi-Yong H., Wei-Jie Z., Xiang-Fu W.;			
RT	"Cloning and Expression of zot Gene from Vibrio cholerae.";			
RL	Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.			
RN	(4)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-EL TOR 86015 / SEROTYPE O1;			
RA	Kan B., Liu Y.Q., Qi G.M., Gao S.Y.;			
RT	Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.			
RN	(5)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-EL TOR N16961 / SEROTYPE O1;			
RA	MEDLINE-20406833; PubMed-10952301;			
RT	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,			
RA	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,			
RA	Gill S.R., Nelson K.E., Read T.D., Tettein H., Richardson D.,			
RA	Ermolaeva M.D., Yamatchevan J., Bass S., Qin H., Dragol I., Sellers P.,			
RA	McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,			
RA	Fraser C.M.;			
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.";			
RL	Nature 406:477-483(2000).			
RN	(6)			
RP	CHARACTERIZATION.			
RA	MEDLINE-91271365; PubMed-2052603;			
RA	Fasano A., Baudry B., Pumphlin D.W., Wasserman S.S., Tall B.D.,			

RA Kelley J.M., Kaper J.B.;
 RT "Vibrio cholerae produces a second enterotoxin, which affects
 RL intestinal tight junctions.";
 CC Proc. Natl. Acad. Sci. U.S.A. 88:5242-5246(1991).
 CC -1- FUNCTION: INCREASES THE PERMEABILITY OF THE SMALL INTESTINE MUCOSA
 CC BY AFFECTING THE STRUCTURE OF INTERCELLULAR TIGHT JUNCTIONS
 CC (ZONULA OCCUDENS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M83563; AAA27582.1; -
 CC DR EMBL: AF175708; AAD51358.1; -
 CC DR EMBL: AF123049; AAD26854.1; -
 CC DR EMBL: AF220606; AAF29547.1; -
 CC DR EMBL: AE004224; AAF94615.1; -
 CC DR PIR: A43864; A43864.
 CC DR TIGR: VC1438; -
 CC KM Enterotoxin; Toxin.
 CC FT VARIANT 45 M -> I (IN STRAIN 569B).
 CC FT VARIANT 100 V -> A (IN STRAINS 569B AND 86015).
 CC FT VARIANT 272 V -> A (IN STRAIN 569B).
 CC FT VARIANT 281 V -> A (IN STRAIN 569B).
 CC FT VARIANT 349 A -> S (IN STRAIN 86015).
 CC FT VARIANT 381 K -> R (IN STRAIN 86015).
 CC FT CONFLICT 386 IKTEENDKCKLSIF -> VKKEKESITKSL (IN REF.
 CC FT
 CC SQ SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;
 CC
 CC Query Match 93.3%; Score 42; DB 1; Length 399;
 CC Best Local Similarity 87.5%; Pred. No. 0.56;
 CC Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 GRVCVODG 8
 CC DB 291 GRICVODG 298
 CC
 CC RESULT 2
 CC DPOL_RCMVM STANDARD; PRT; 1120 AA.
 CC ID DPOL_RCMVM
 CC AC 083428;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 15-DEC-1998 (Rel. 37, Last annotation update)
 CC DE DNA POLYMERASE (EC 2.7.7.7).
 CC GN ULS4.
 CC OS Rat cytomegalovirus (strain Maastricht).
 CC CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC CC Betaherpesvirinae; Muromegalovirus.
 CC CC NCBI_TaxID=79700;
 CC CC (1)
 CC CC SEQUENCE FROM N.A.
 CC CC MEDLINE=96335691; PubMed=8757999;
 CC CC Beuken E., Slobbe R., Bruggeman C.A., Vink C.;
 CC CC "Cloning and sequence analysis of the genes encoding DNA polymerase,
 CC CC glycoprotein B, ICP18.5 and major DNA-binding protein of rat
 CC CC cytomegalovirus.";
 CC CC J. Gen. Virol. 77:1559-1562(1996).
 CC CC -1- CATALYTIC ACTIVITY: N DPOXYNUCLEOSIDE TRIPHOSPHATE =
 CC CC N PROPHOSPHATE + DNA(N).
 CC CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 CC CC -----
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC -----

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: U50550; AAC56433.1; -
 CC DR InterPro: IPR002064; -
 CC DR Pfam: PF00136; DNA_POL_B; 1.
 CC DR PRINTS: PR00106; DNA_POLB.
 CC DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
 CC KW Transferrase; DNA-directed DNA polymerase; DNA replication;
 CC KW DNA-binding; Nuclear protein.
 CC SQ SEQUENCE 1120 AA; 124853 MW; 12FFAE95EF54FC4 CRC64;
 CC
 CC Query Match 82.28; Score 37; DB 1; Length 1120;
 CC Best Local Similarity 75.0%; Pred. No. 14;
 CC Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 GRVCVODG 8
 CC DB 493 GRVCVODG 500
 CC
 CC RESULT 3
 CC ID NUSG_CGDAB STANDARD; PRT; 70 AA.
 CC AC P36262;
 CC DT 01-JUN-1994 (Rel. 29, Created)
 CC DT 01-JUN-1994 (Rel. 29, Last sequence update)
 CC DT 01-NOV-1995 (Rel. 32, Last annotation update)
 CC DE TRANSCRIPTION ANTI-TERMINATION PROTEIN NUSG (FRAGMENT).
 CC GN NUSG.
 CC OS Citrus greening disease-associated bacterium-like organism.
 CC CC Bacteria; Proteobacteria; gamma subdivision.
 CC CC NCBI_TaxID=2707;
 CC CC (1)
 CC CC SEQUENCE FROM N.A.
 CC CC Vilcheanoux S., Garnier M., Laigret F., Renaudin J., Bove J.M.;
 CC CC Submitted (XX-1992) to the EMBL/GenBank/DBJ databases.
 CC CC -1- FUNCTION: INFLUENCES TRANSCRIPTION TERMINATION AND
 CC CC ANTI-TERMINATION. ACTS AS A COMPONENT OF THE TRANSCRIPTION COMPLEX,
 CC CC AND INTERACTS WITH THE TERMINATION FACTOR RHO AND RNA POLYMERASE
 CC CC (BY SIMILARITY).
 CC CC -1- SIMILARITY: BELONGS TO THE NUSG FAMILY.
 CC CC -----
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC -----
 CC CC EMBL: M94319; AAA23104.1; -
 CC CC InterPro: IPR001062; -
 CC CC DR PROSITE: PS01014; NUSG; 1.
 CC CC KW Transcription termination.
 CC CC NON_TER 1
 CC SQ SEQUENCE 70 AA; 7749 MW; 563C62DB9A0C485 CRC64;
 CC
 CC Query Match 75.68; Score 34; DB 1; Length 70;
 CC Best Local Similarity 85.7%; Pred. No. 3.6;
 CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 2 RVCVODG 8
 CC DB 21 RVCVSDG 27
 CC
 CC RESULT 4
 CC SWE_CHICK

ID SM3E.CHICK STANDARD: PRT: 785 AA.
AC 0A4237: 090666;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE SEMAPHORIN 3E PRECURSOR (COLLAPLIN-5) (COLL-5).
GN SEMA3E OR COLL5.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=97470887; PubMed=9333347;
RA Felner L., Koppel A.M., Kobayashi H., Raper J.A.;
RT "Secreted chick semaphorins bind recombinant neuropilin with similar
affinities but bind different subsets of neurons in situ.";
RL Neuron 19:539-545(1997).
RN [2]
RP SEQUENCE OF 244-543 FROM N.A.
RX MEDLINE=9539269; PubMed=7605628;
RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;
RT "A family of molecules related to collapsin in the embryonic chick
nervous system.";
RL Neuron 14:1131-1140(1995).
CC -1- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
CC NEURONAL POPULATIONS. BINDS TO NEURAPILIN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: COLLAPLIN-1, -2, -3, AND -5 BIND TO
OVERLAPPING BUT DISTINCT AXON TRACTS.
CC -1- DOMAIN: STRONG BINDING TO NEURAPILIN IS MEDIATED BY THE CARBOXY
THIRD OF THE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF022947; AAB80952.1; -;
DR EMBL: U28243; AAA86899.1; -;
DR InterPro: IPR000292; -;
DR InterPro: IPR001627; -;
DR Pfam: PFO1403; Sema.1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 25
FT CHAIN 26 785
FT DOMAIN 244 544
FT DOMAIN 651 740
FT DOMAIN 741 780
FT DISULFID 658 733
FT CARBOHYD 48 48
FT CARBOHYD 130 130
FT CARBOHYD 600 600
FT CONFLICT 246 246
FT CONFLICT 248 248
FT CONFLICT 250 250
FT SEQUENCE 785 AA; 90978 MW; E551BEF717630632 CRC64;
Query Match 75.6%; Score 34; DB 1; Length 785;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 GRVCVOD 7

DB 271 GRVCVND 277
|||||
RESULT 5
ID YRS7 CAEEL STANDARD: PRT: 796 AA.
AC 010003;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 90.8 KDA PROTEIN T05H10.7 IN CHROMOSOME II.
GN T05H10.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Lightning J., Thomas K.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z47811; CA887788.1; -;
DR EMBL: Z47812; CA887788.1; JOINED.
DR EMBL: Z47812; CA887796.1; -;
DR Wormpep: T05H10.7; CE03637.
KW Hypothetical protein.
FT DOMAIN 38 42
FT DOMAIN 67 70
FT DOMAIN 524 527
FT SEQUENCE 796 AA; 90831 MW; 7BDF8E0A4D2AA9F1 CRC64;
Query Match 75.6%; Score 34; DB 1; Length 796;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 GRVCVODG 8
|||||
DB 198 GRKCVSDG 205
RESULT 6
ID DPOL_HSVT1 STANDARD: PRT: 1171 AA.
AC 09YUS3;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DNA POLYMERASE (EC 2.7.7.7).
GN DPOL.
OS Herpesvirus tupaia (Strain 1) (THV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=132677;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98094630; PubMed=9880021;
RA Springfield C., Tildona C.A., Kehm R., Bahr U., Darai G.;
RT "Identification and characterization of the Tupaia herpesvirus DNA
polymerase gene.";
RL J. Gen. Virol. 79:3049-3053(1998).
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -

CC N PYROPHOSPHATE + DNA(N).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF074327; AAD08667.1; -
 CC Interpro: IPR002064; -
 CC Pfam: PF00136; DNA_POL_B; 2.
 CC PRINTS: PR00106; DNAPOLB.
 CC PROSITE: PS00116; DNA_POLYMERASE_B; 1.
 CC Transferrase; DNA-directed DNA polymerase; DNA replication;
 CC DNA-binding; Nuclear protein
 CC SEQUENCE 1171 AA; 128622 MW; CDC0480FEACFC7EC CRC64;
 CC
 CC Query Match
 CC Best Local Similarity 75.6%; Score 34; DB 1; Length 1171;
 CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC Oy 1 GRVCVOD 7
 CC Db 515 GRVCVOD 521
 CC
 CC RESULT 7
 CC DPOL_HSVT2
 CC ID DPOL_HSVT2 STANDARD; PRT; 1171 AA.
 CC AC Q9Y0S2;
 CC DT 01-OCT-2000 (Rel. 40; Created)
 CC DT 01-OCT-2000 (Rel. 40; Last sequence update)
 CC DE 01-OCT-2000 (Rel. 40; Last annotation update)
 CC DE DNA POLYMERASE (EC 2.7.7.7).
 CC GN DPOL.
 CC OS Herpesvirus tupaia (Strain 2) (THV-2).
 CC OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC OX NCBI_TaxID=132678;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RA MEDLINE=99094630; PubMed=9880021;
 CC RA Springfield C., Tidona C.A., Kelm R., Bahr U., Darai G.;
 CC RT "Identification and characterization of the Tupaia herpesvirus DNA
 CC RT polymerase gene.";
 CC RL J. Gen. Virol. 79:3049-3053(1998).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RA MEDLINE=99319892; PubMed=10392721;
 CC RA Bahr U., Springfield C., Tidona C.A., Darai G.;
 CC RT "Structural organization of a conserved gene cluster of Tupaia
 CC RT herpesvirus encoding the DNA polymerase, glycoprotein B, a probable
 CC RT processing and transport protein, and the major DNA binding protein.";
 CC RL Virus Res 60:123-136(1999).
 CC RN [3]
 CC RP CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
 CC N PROPHOSPHATE + DNA(N).
 CC CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF074328; AAD08667.1; -

DR EMBL: AF084543; AAD2936.1; -
 DR Interpro: IPR002064; -
 DR Pfam: PF00136; DNA_POL_B; 2.
 DR PRINTS: PR00106; DNAPOLB.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
 DR Transferrase; DNA-directed DNA polymerase; DNA replication;
 DR DNA-binding; Nuclear protein.
 DR SEQUENCE 1171 AA; 128590 MW; D2D64897FDE570E8 CRC64;
 DR
 DR Query Match
 DR Best Local Similarity 75.6%; Score 34; DB 1; Length 1171;
 DR Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 DR
 DR Oy 1 GRVCVOD 7
 DR Db 515 GRVCVOD 521
 DR
 DR RESULT 8
 DR ASPG_FLAME
 DR ID ASPG_FLAME STANDARD; PRT; 340 AA.
 DR AC 047898;
 DR DT 01-NOV-1997 (Rel. 35; Created)
 DR DT 01-NOV-1997 (Rel. 35; Last sequence update)
 DR DE 30-MAY-2000 (Rel. 39; Last annotation update)
 DR DE N4-(BETA-N-ACETYLGLUCOSAMINYL)-L-ASPARAGINASE PRECURSOR (EC 3.5.1.26)
 DR DE (GLYCOSYLASPARAGINASE) (ASPARTYLGLUCOSAMINIDASE) (N4-(N-ACETYL-BETA-
 DR DE FLAVOBACTERIUM MENINGOSEPTICUM) (AGA).
 DR OC Bacteria; CPB group; Flavobacteria; Flavobacteriaceae;
 DR Chryseobacterium.
 DR OX NCBI_TaxID=238;
 DR RN [1]
 DR RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 DR RC STRAIN=ELDER / ATCC 33958;
 DR RX MEDLINE=95142655; PubMed=7840643;
 DR RA Tarentino A.L., Quinones G., Hauer C.R., Changchien L.-M.,
 DR RA Plummer T.H. Jr.;
 DR RT "Molecular cloning and sequence analysis of Flavobacterium
 DR RT meningosepticum glycosylasparaginase: a single gene encodes the alpha
 DR RL Arch. Biochem. Biophys. 316:399-406(1995).
 DR RN [2]
 DR RP SEQUENCE OF 46-59 AND 197-211.
 DR RX MEDLINE=94071939; PubMed=8250923;
 DR RA Tarentino A.L., Plummer T.H. Jr.;
 DR RT "The first demonstration of a prokaryotic glycosylasparaginase.";
 DR RL Biochem. Biophys. Res. Commun. 197:179-186(1993).
 DR RN [3]
 DR RP X-RAY CRYSTALLOGRAPHY (2.32 ANGSTROMS).
 DR RX MEDLINE=98200483; PubMed=9541410;
 DR RA Xuan J., Tarentino A.L., Grimwood B.G., Plummer T.H. Jr., Cui T.,
 DR RA Guan C., van Roey P.;
 DR RT "Crystal structure of glycosylasparaginase from Flavobacterium
 DR RL meningosepticum.";
 DR RL Protein Sci. 7:774-781(1998).
 CC -1- FUNCTION: CLEAVES THE GLICNAc-ASN BOND WHICH JOINS OLIGOSACCHARIDES
 CC TO THE PEPTIDE OF ASPARAGINE-LINKED GLYCOPROTEINS. REQUIRES THAT
 CC THE GLYCOSYLATED ASPARAGINE MOIETY IS NOT SUBSTITUTED ON
 CC ITS AMINO (R1) AND CARBOXYL (R2) TERMINUS.
 CC -1- CATALYTIC ACTIVITY: N4-(BETA-N-ACETYL-D-GLUCOSAMINYL)-L-ASPARAGINE
 CC + H2O = N-ACETYL-BETA-GLUCOSAMINYLAMINE + L-ASPARATE.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE ASPARAGINASE 2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>


```

CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: U08028: AAA6868.1; -.
DR PDB: 1AYI: 29-APR-98.
DR MEROPS: T02.001; -.
DR InterPro: IPR000246; -.
DR Pfam: PF01112: Asparaginase_2; 1.
KM Signal: Hydrolase; Periplasmic; 3D-structure.
FT SIGNAL 1 45
FT CHAIN 196 196 GLYCOSYLASPARAGINASE, ALPHA CHAIN.
FT CHAIN 197 340 GLYCOSYLASPARAGINASE, BETA CHAIN.
FT DISULFID 113 126
FT DISULFID 213 277 OR 315.
FT ACT SITE 197 197 BY SIMILARITY.
SQ SEQUENCE 340 AA: 37262 MM: 4C56E5061B4E53D7 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 340;
Best Local Similarity 62.5%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRVCVODG 8
1 1111
Db 312 GAVCVDG 319

RESULT 9
GLTD_ECOLI STANDARD; PRT; 471 AA.
AC P09832;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GLUTAMATE SYNTHASE [NADPH SMALL CHAIN (EC 1.4.1.13) (GLUTAMATE
DE SYNTHASE BETA SUBUNIT) (NADPH-GOGAT) (GLTS BETA CHAIN)].
GN GLTD OR ASPB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxId=562;
RN [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-K12;
RX MEDLINE=88152432; PubMed=3326786;
RA Oliver G., Gosec G., Sanchez-Pescador R., Lozoya E., Ku L.M.,
RA Flores N., Becerril B., Valle F., Bolivar F.;
RT "Determination of the nucleotide sequence for the glutamate synthase
RT structural genes of Escherichia coli K-12.";
RL Gene 60:1-11(1987).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97443975; PubMed=9298646;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.U.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RN SEQUENCE OF 1-12.
RC STRAIN-K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.O., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
RN [4]
RN DISCUSSION OF SEQUENCE.
RP MEDLINE=89098858; PubMed=2643092;
RX Gosec G., Merino E., Recillas F., Oliver G., Becerril B., Bolivar F.;
RT "Amino acid sequence analysis of the glutamate synthase enzyme from
RT Escherichia coli K-12.";

```

```

RL Protein Seq. Data Anal. 2:9-16(1989).
CC -1- CATALYTIC ACTIVITY: 2 L-GLUTAMATE + NADP(+) -> L-GLUTAMINE +
CC 2-OXOGLOUTARATE + NADPH.
CC -1- COFACTOR: BINDS A 4FE-4S CLUSTER.
CC -1- PATHWAY: NITROGEN METABOLISM, GLUTAMATE BIOSYNTHESIS.
CC THE CATALYZED REACTION BRINGS TOGETHER THE NITROGEN AND
CC CARBON METABOLISM.
CC -1- SUBUNIT: AGGREGATE OF 4 CATALYTIC ACTIVE HETERODIMERS,
CC CONSISTING OF A LARGE AND A SMALL SUBUNIT.
CC -1- MISCELLANEOUS: GLUTAMINE BINDS TO THE LARGE SUBUNIT AND TRANSFERS
CC THE AMIDO GROUP TO 2-OXO-GLUTAMATE THAT APPARENTLY BINDS TO THE
CC SMALL SUBUNIT.
CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
CC -1- SIMILARITY: STRONG TO E. COLI AEGA AND YGFT.
CC -----
CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: M18747: AAA23905.1; -.
DR EMBL: U18997: AAA58015.1; -.
DR EMBL: AE000400: AAC76245.1; -.
DR PIR: B29617: B29617.
DR SWISS-2DPAGE: P09832: COLI.
DR EC02DBASE: F050.4: 6TH EDITION.
DR Ecocore: EG10404: g1td.
DR InterPro: IPR000759; -.
DR PRINTS: PR00419; ADXROTASE.
DE Oxidoreductase; Glutamate biosynthesis; Iron-sulfur; 4Fe-4S; NADP.
FT FT INTL_MET 0 0
FT METAL 93 93 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 97 97 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 103 103 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 107 107 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT CONFLICT 37 50 GOAQAQADRLCSG -> ARKRRLTACRAA (IN
FT REF. 1).
FT CONFLICT 122 122 E -> K (IN REF. 1).
FT CONFLICT 173 173 V -> C (IN REF. 1).
FT CONFLICT 256 269 VYALPFLIANTKQ -> CYORCRSSSTPENS (IN
FT REF. 1).
FT CONFLICT 311 312 KH -> ND (IN REF. 1).
FT CONFLICT 375 399 GRRRAETVAGSEHIVPADAVIMAFG -> ASPRGDCRCFRFT
FT YRTGRCGDHGW (IN REF. 1).
SQ SEQUENCE 471 AA: 51884 MM: C4070F7509C561A8 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 471;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVCVOD 7
111111
Db 94 GRCVOD 100

RESULT 10
TRPE_SYNY3
ID TRPE_SYNY3 STANDARD; PRT; 508 AA.
AC P20170; P20168;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE ANTHRANILATE SYNTHASE COMPONENT 1 (EC 4.1.3.27).
GN TRPE OR SLR0738.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxId=1148;
RN [1]

```

RP SEQUENCE FROM N.A.
 RA MEDLINE-97061201; PubMed-8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nario K.,
 RA Okumura S., Shimizu S., Takeuchi C., Wada T., Watanabe A.,
 RA Yamada M., Yasuda M., Tabata S.,
 RA "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 RN 121
 RP SEQUENCE OF 1-408 FROM N.A.
 RX MEDLINE-89034300; PubMed-311423;
 RA Reilly P., Hulmes J.D., Pan Y.C.E., Nelson N.,
 RT "Molecular cloning and sequencing of the psad gene encoding subunit
 RT II of photosystem I from the cyanobacterium, Synechocystis sp. PCC
 RL 6803.";
 RL J. Biol. Chem. 263:17658-17662(1988).
 CC -1- CATALYTIC ACTIVITY: CHORISMATE + L-GLUTAMINE - ANTHRANILATE +
 CC PYRUVATE + L-GLUTAMATE.
 CC -1- PATHWAY: FIRST STEP IN BIOSYNTHESIS OF TRYPTOPHAN.
 CC -1- SUBUNIT: TETRAMER OF TWO COMPONENTS I AND TWO COMPONENTS II (BY
 CC SIMILARITY).
 CC -1- MISCELLANEOUS: COMPONENT I CATALYZES THE FORMATION OF ANTHRANILATE
 CC USING AMMONIA RATHER THAN GLUTAMINE, WHEREAS COMPONENT II PROVIDES
 CC GLUTAMINE AMIDOTRANSFERASE ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE ANTHRANILATE SYNTHASE COMPONENT I
 CC FAMILY.
 CC -1- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 CC FRAMESHIFT IN POSITION 182.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D90899; BAA16689.1; -
 DR EMBL; J04195; AAA88626.1; -
 DR EMBL; J04195; AAA88627.1; ALT_FRAME.
 DR PIR; B32124; B32124.
 DR PIR; C32124; C32124.
 DR InterPro: IPR000350; -
 DR Pfam: PF00435; Chorismate-bind; 1.
 DR PRINTS: PR00095; ANTSNTTHASEI.
 DR Trpophan biosynthesis; Lyase.
 KM Tyrosine biosynthesis; Lyase.
 SQ SEQUENCE 508 AA; 57147 MW; A3308E9A7B8A1FC1 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 508;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GRVCVQ 6
 DB 363 GRVCVQ 368

RESULT 11
 GLPK_MOUSE
 ID GLPK_MOUSE STANDARD; PRT; 524 AA.
 AC 064516;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE)
 DE (GLYCEROKINASE) (GK).
 GN GK OR GYK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RX STRAIN-BALB/C;
 RA MEDLINE-97038697; PubMed-8884278;
 RA Hug A.H., Lovell R.S., Sampson M.J., Decker W.K., Dinulos M.B.,
 RA Distche C.M., Craig W.J.,
 RT "Isolation, mapping, and functional expression of the mouse x
 RT chromosome glycerol kinase gene.";
 RL Genomics 36:530-534(1996).
 CC -1- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND
 CC METABOLISM (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + GLYCEROL - ADP + GLYCEROL 3-PHOSPHATE.
 CC -1- PATHWAY: RATE-LIMITING STEP IN GLYCEROL UTILIZATION.
 CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL SURFACE OR
 CC CYTOSOL (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCOKINASE /
 CC GLYCEROKINASE / XYLOKINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U48403; AAC52824.1; -
 DR HSSP; P08859; 1GB.
 DR MGD; MGI:106594; GYK.
 DR InterPro: IPR000577; -
 DR Pfam: PF00370; FGGY; 1.
 DR PROSITE; PS00445; FGGY_KINASES_2; 1.
 DR PROSITE; PS00933; FGGY_KINASES_1; 1.
 KM Glyceral metabolism; Transferase; Kinase; ATP-binding.
 FT NP_BIND 167 179 ATP (PROBABLE).
 SQ SEQUENCE 524 AA; 57458 MW; C3E1C56887683B55 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 524;
 Best Local Similarity 62.5%; Pred. No. 39;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRVCVQDG 8
 DB 266 GOMCPODG 273

RESULT 12
 GLPK_RAT
 ID GLPK_RAT STANDARD; PRT; 524 AA.
 AC 063060;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE)
 DE (GLYCEROKINASE) (GK) (ATP-STIMULATED GLUCOCORTICOID-RECEPTOR
 DE TRANSLATION PROMOTER) (ASTP).
 GN GK.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN 11
 RP SEQUENCE FROM N.A.
 RX STRAIN-SPRAGUE-DAWLEY; TISSUE-Liver;
 RA MEDLINE-93312338; PubMed-8323560;
 RA Okamoto K., Hirano H., Iwashashi F.,
 RT "Molecular cloning of rat liver glucocorticoid-receptor translocation
 RT promoter.";
 RL Biochem. Biophys. Res. Commun. 193:848-854(1993).
 CC -1- FUNCTION: INCREASES THE BINDING OF ACTIVATED GLUCOCORTICOID-

```
CC RECEPTOR TO NUCLEI IN THE PRESENCE OF ATP.
CC -1- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND
CC METABOLISM (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + GLYCEROL = ADP + GLYCEROL 3-PHOSPHATE.
CC -1- PATHWAY: RATE-LIMITING STEP IN GLYCEROL UTILIZATION.
CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL SURFACE OR
CC CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCOKINASE /
CC GLYCEROKINASE / XYLOUKINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR EMBL: D16102; BAA03677.1; -

DR HSP: P08859; IGLB.

DR InterPro: IPR000577; -

DR Pfam: PF00370; FGGY; 1.

DR PROSITE: PS00933; FGGY_KINASES_1; 1.

DR PROSITE: PS00445; FGGY_KINASES_2; 1.

KM Glycerol metabolism; Transferase; Kinase; ATP-binding.

FT NP_BIND 167 179 ATP (PROBABLE).

SO SEQUENCE 524 AA; 57477 MW; A620C296D32B684 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 524;
Best Local Similarity 62.5%; Pred. No. 39;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVCVODG 8
|:| | | |

Db 266 GOMCFODG 273

RESULT 13
YGF1_ECOLI STANDARD: PRT; 639 AA.

AC 046820;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HYPOTHETICAL 69.1 KDA PROTEIN IN KDOI-LYSS INTERGENIC REGION.
GN YGF1.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-K12 / MG1655;
RX MEDLINE-97426617; Pubmed-9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -1- COFACTOR: BINDS TWO 4FE-4S CLUSTERS (POTENTIAL).
CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC 'BACTERIAL-TYPE' 4FE-4S FERRODOXINS.
CC -1- SIMILARITY: STRONG, TO E. COLI GLUTAMATE SYNTHASE (NADPH) SMALL
CC CHAIN (GLTD) AND TO E. COLI AEGA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```
CC or send an email to license@isb-sib.ch).
```

DR EMBL: U28375; AAA83068.1; ALT_INIT.

DR EMBL: AE000372; AAC75925.1; ALT_INIT.

DR HSSP: P00198; IFDN.

DR EcGene: EG13070; YGF1.

DR InterPro: IPR000759; -

DR InterPro: IPR001450; -

DR Pfam: PF00037; Ier4; 1.

DR PRINTS: PR00353; 4FE4SFDOXIN.

DR PRINTS: PR00419; ADXRDASE.

DR PROSITE: PS00198; 4FE4S_FERRODOXIN; 1.

KM Hydrothermal protein; Oxidoreductase; Iron-sulfur; 4Fe-4S.

FT METAL 87 87 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).

FT METAL 90 90 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).

FT METAL 93 93 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).

FT METAL 97 97 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).

FT METAL 258 258 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).

FT METAL 262 262 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).

FT METAL 268 268 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).

FT METAL 272 272 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).

SO SEQUENCE 639 AA; 69088 MW; 3B2B96621835655C CRC64;

Query Match 73.3%; Score 33; DB 1; Length 639;
Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVCVOD 7
| | | | |

Db 259 GRVCPOD 265

RESULT 14
AEGA_ECOLI STANDARD: PRT; 659 AA.

AC P37127; P76560; P76970;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE AEGA PROTEIN.
GN AEGA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-K12 / MC4100;
RX MEDLINE-97113461; Pubmed-8955321;
RA Cavicholi R., Kolesnikow T., Gunsalus R.P.;
RT "Characterization of the aega locus of Escherichia coli: control of
RT gene expression in response to anaerobiosis and nitrate.";
RL J. Bacteriol. 178:6968-6974(1996).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-K12 / MG1655;
RX MEDLINE-97426617; Pubmed-9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN-K12;
RC Alpha H., Baba T., Fujita K., Hayashi K., Honjo A., Horuchi T.,
RA Ikemoto K., Inada T., Isono K., Itoh T., Kanai K.,
RA Kasai H., Kashimoto K., Kim S., Kimura S., Kitagawa M.,
RA Kitakawa M., Makino K., Masuda S., Miki T., Mizobuchi K., Mori H.,
RA Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sempel G., Seki Y., Tagami H., Takemoto K., Wada C.,

```

RT      "Identification and structural characterization of the 75-kDa rabbit
RT      zona pellucida protein."
CC      J. Biol. Chem. 268:12412-12417(1993).
CC      -I- FUNCTION: ZP2 FORMS WITH ZP1 AND ZP3 THE ZONA PELLUCIDA, IN
CC      WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC      ZP2 ACTS AS A SECONDARY SPERM RECEPTOR (BY SIMILARITY).
CC      -I- SUPRACELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXTRACELLULAR
CC      MATRIX.
CC      -I- PTM: PROPEPTIDITICALLY CLEAVED AFTER FERTILIZATION, AND THIS
CC      MODIFICATION, ALONG WITH PRESUMED CHANGES IN ZP3 MAY PLAY AN
CC      IMPORTANT ROLE IN THE POSTFERTILIZATION BLOCK TO POLYSPERMY.
CC      -I- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES
CC      (BY SIMILARITY).
CC      -----
CC      -I- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outpost at
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC      or send an email to license@isb-slb.ch).
CC      -----
DR      EMBL: L12167; AAA31502.1; .
DR      InterPro: IPR001507; .
DR      Pfam: PF00100; zona_pellucida; 1.
DR      ProSITE: PS00682; ZP_DOMAIN; 1.
KW      Glycoprotein; Sulfatation; Sperm; Receptor; Transmembrane;
KM      Extracellular matrix.
FT      NON_TER          1
FT      CHAIN            1
FT      DOMAIN           <1
FT      TRANSMEM        637
FT      TRANSMEM        657
FT      DOMAIN           318
FT      DOMAIN           585
FT      CARBOHYD         38
FT      CARBOHYD         73
FT      CARBOHYD        126
FT      CARBOHYD        171
FT      CARBOHYD        217
FT      CARBOHYD        241
FT      CARBOHYD        613
SQ      SEQUENCE        666 AA; 73644 MW; DDC8E2BA2D21020B CRC64;
Query Match       73.3%; Score 33; DB 1; Length 666;
Best Local Similarity 62.5%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY      1 GRVGVDS 8
DB      317 GELCVDG 324

```

Search completed: June 13, 2001, 14:21:39
Job time: 799 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:20:21 ; Search time 225.85 Seconds
(without alignments)
4.152 Million cell updates/sec

Title: PCT-US01-05825A-2
Sequence: 1 GRVCVQDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP.TREMBL.15:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.podent:*
13: sp.unclassified:*
14: sp.vertebrate:*
15: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	93.3	323	2	09L8F5
2	42	93.3	399	2	09R3V6
3	42	93.3	399	2	09L706
4	38	84.4	217	4	014964
5	36	80.0	516	2	084333
6	36	80.0	516	2	09PK65
7	36	80.0	611	2	069514
8	36	80.0	2337	13	09T8G7
9	35	77.8	475	10	081325
10	35	77.8	549	11	09W064
11	35	77.8	984	9	080102
12	34	75.6	782	4	09NS98
13	34	75.6	1048	14	069390
14	33	73.3	185	2	09K852
15	33	73.3	196	2	084353
16	33	73.3	196	2	09PK45
17	33	73.3	437	2	09Z5V6
18	33	73.3	440	5	076463
19	33	73.3	471	2	09K8J0

20	33	73.3	477	2	09S457	09S457 pseudomonas
21	33	73.3	494	2	09P811	09P811 xyloella fas
22	33	73.3	551	11	061192	061192 mus musculu
23	33	73.3	582	5	022811	022811 caenorhabdi
24	33	73.3	870	3	014424	014424 coccidioid
25	33	73.3	1007	14	09P817	09P817 porcine cyt
26	33	73.3	1007	14	09P816	09P816 porcine cyt
27	33	73.3	1007	14	09P815	09P815 porcine cyt
28	33	73.3	1007	14	09P813	09P813 porcine cyt
29	33	73.3	1013	14	092827	092827 human herpe
30	33	73.3	1551	6	09MZP4	09MZP4 canis fami
31	32	71.1	143	10	052062	052062 enterobacte
32	32	71.1	143	10	09S842	09S842 arabidopsi
33	32	71.1	224	2	09Z6S3	09Z6S3 chlamydia p
34	32	71.1	259	1	033163	033163 methanosarc
35	32	71.1	267	5	09N9U5	09N9U5 leishmania
36	32	71.1	326	10	040316	040316 medicago sa
37	32	71.1	326	10	09ZPK2	09ZPK2 pisum sativ
38	32	71.1	326	10	09SDP2	09SDP2 glycine max
39	32	71.1	341	5	09VD06	09VD06 drosophila
40	32	71.1	342	2	09Z4N6	09Z4N6 pasteurilla
41	32	71.1	356	11	062559	062559 murinae gen
42	32	71.1	413	2	008340	008340 rhodobacter
43	32	71.1	417	10	049465	049465 arabidopsi
44	32	71.1	419	5	024548	024548 drosophila
45	32	71.1	419	5	09VBAS	09VBAS drosophila

ALIGNMENTS

RESULT 1	09L8F5	PRELIMINARY;	PRT;	323 AA.
ID	09L8F5			
AC	09L8F5			
DT	01-OCT-2000 (TREMBlrel. 15, Created)			
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)			
DT	01-OCT-2000 (TREMBlrel. 15, Last annotation update)			
DE	ZOT (FRAGMENT).			
GN	ZOT.			
OS	Vibrio mimicus.			
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.			
OX	NCBI_TaxID=674;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=PT5;			
RX	MEDLINE=20143766; PubMed=10678967;			
RA	Boyd E.F., Moyer K.E., Shi L., Waldor M.K.;			
RT	"Infectious CTXphi and the vibrio pathogenicity island prophage in			
RT	vibrio mimicus: evidence for recent horizontal transfer between V.			
RT	mimicus and V. cholerae."			
RL	Infect. Immun. 68:1507-1513(2000).			
DR	EMBL; AF207857; AAF40142.1; -.			
FT	NON_TER	1		
FT	NON_TER	1		
SO	SEQUENCE	323 AA;	323 MM;	01C12DAE9B873C3B CRC64;
Query Match	93.3%;	Score 42;	DB 2;	Length 323;
Best Local Similarity	87.5%;	Pred. No. 1.2;		
Matches	7;	Conservative	1;	Mismatches 0; Indels 0;
OY	1 GRVCVQDG 8			
Db	240 GRVCVQDG 247			
RESULT 2	09R3V6	PRELIMINARY;	PRT;	399 AA.
ID	09R3V6			
AC	09R3V6			
DT	01-MAY-2000 (TREMBlrel. 13, Created)			
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)			

DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE ZONULAR OCCLUDENS TOXIN (ZONA OCCLUDENS TOXIN).
 GN ZOF OR VC1458.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KNH002;
 RA Shin H.J., Park Y.C., Kim Y.C.;
 RT "Cloning and nucleotide sequence analysis of the virulence gene
 RT cassette from Vibrio cholerae KNH002 isolated in Korea.";
 RL Misimurha Hoji 35:205-210(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O139-TOR OGAWA;
 RA Zhi-Yong H., Wei-Jie Z., Xiang-Fu W.;
 RT "Cloning and Expression of zot Gene from Vibrio cholerae.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EL TOR N16961 / SEROTYPE O1;
 RA MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McLeod S.L., Utterback T., Fleischmann R.D., Nieman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 DR EMBL: AF15708; AAD51358.1; -;
 DR EMBL: AF15708; AAD51358.1; -;
 DR EMBL: AF004224; AAF94615.1; -;
 DR TIGR: VC1458; -;
 SQ SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 399;
 Best Local Similarity 87.5%; Pred. No. 1.5;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
 DB 291 GRVCVODG 298

RESULT 3
 ID 09L706 PRELIMINARY; PRT; 399 AA.
 AC 09L706;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 GN ZOT.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-86015;
 RA Kan B., Liu Y.Q., Qi G.M., Gao S.Y.;
 RT "Vibrio cholerae nct-cryphi whole genome, include rstr(RstR),
 RT rtsA(RtsA), rtsB(RstB), cep(Cep), orfU(OrfU), ace(Ace) and zot(Zot)
 RT genes.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF220606; AAF29547.1; -;
 SQ SEQUENCE 399 AA; 44990 MW; CF6A3DBCC9E23EE1 CRC64;

Query Match 93.3%; Score 42; DB 2; Length 399;
 Best Local Similarity 87.5%; Pred. No. 1.5;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
 DB 291 GRVCVODG 298

RESULT 4
 ID 014964 PRELIMINARY; PRT; 217 AA.
 AC 014964;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE RAB-RELATED GTP-BINDING PROTEIN (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=97237046; PubMed=9119394;
 RA Stankovic T., Byrd P.J., Cooper P.R., McConville C.M., Munroe D.J.,
 RA Riley J.H., Watts G.D., Ambrose R., McGuire G., Smith A.D.,
 RA Sutcliffe A., Mills T., Taylor A.M.R.;
 RT "Construction of a transcription map around the gene for ataxia
 RT telangiectasia: identification of at least four novel genes.";
 RL Genomics 40:267-276(1997).
 DR EMBL: X99962; CAA68227.1; -;
 DR HSSP: P05713; 3RAB.
 DR INTERPRO: IPR001806; -;
 DR PFAM: PF00071; Ras1.1.
 DR PRINTS: PR00449; RASTRSEFRNG.
 FT NON_TER 1 1
 FT 217 217
 SQ SEQUENCE 217 AA; 24869 MW; AA9A29731F42B8F5 CRC64;

Query Match 84.4%; Score 38; DB 4; Length 217;
 Best Local Similarity 62.5%; Pred. No. 5;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
 DB 182 GEICIDQG 189

RESULT 5
 ID 084333 PRELIMINARY; PRT; 516 AA.
 AC 084333;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, last annotation update)
 GN EXODOXIRIBONUCLEASE VII.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-D/WH-3/CX;
 RA MEDLINE=9900809; PubMed=9784136;
 RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 RT Chlamydia trachomatis.";
 RL Science 282:754-759(1998).
 DR EMBL: AE001305; AAC67922.1; -;
 SQ SEQUENCE 516 AA; 58621 MW; BECE3F0FB66A5C0E CRC64;

Query Match 80.0%; Score 36; DB 2; Length 516;
 Best Local Similarity 85.7%; Pred. No. 28;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVCVODG 8
 DB 493 RVCLODQ 499

RESULT 6
 O9PK65 PRELIMINARY; PRT; 516 AA.

AC O9PK65:
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE EXODEOXYRIBONUCLEASE, LARGE SUBUNIT.
 GN TC0605.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 NCBI_TaxID=83560;
 RX MEDLINE=20150255; PubMed=10684935.
 RC STRAIN=MOPN / NIGC;
 RP SEQUENCE FROM N.A.
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., White O.,
 White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
 Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 Gwinn M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MOPN and Chlamydia
 pneumoniae AR39."
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Read T.D., Brunham R., Shen C., Gill S.R., Heidelberg J.F., White O.,
 Hickey E.K., Peterson J., Uterback T., Berry K.,
 Bass S., Linher K., Weidman J., Khouri H., Craven B., Bowman C.,
 Dodson R., Gwinn M., Nelson W., Deboy R., Kolonay J., McClarty G.,
 Salzberg S.L., Eisen J., Fraser C.M.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB026129; AAF39436.1; -.
 DR TIGR: TC0605; -.
 SO SEQUENCE 516 AA; 58882 MW; FBBF029768026EB1 CRC64;

Query Match 80.0%; Score 36; DB 2; Length 516;
 Best Local Similarity 85.7%; Pred. No. 28;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVCVODG 8
 DB 493 RVCLODQ 499

RESULT 7
 O69514 PRELIMINARY; PRT; 611 AA.

AC O69514:
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DE DNA POLYMERASE III SUBUNIT GAMMA AND TAU.
 GN DNAX.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1169;
 RX SEQUENCE FROM N.A.
 RP Seeger K., Harris D.;

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Parthill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9318700; PubMed=8446027;
 RA Eigmeier K., Honore N., Woods S.A., Caudron B., Cole S.T.;
 RT "Use of an ordered cosmid library to deduce the genomic organization
 of Mycobacterium leprae."
 RL Mol. Microbiol. 7:197-206(1993).
 DR EMBL: AL023596; CAA19155.1; -.
 DR INTERPRO: IPR00862; -.
 SO SEQUENCE 611 AA; 65535 MW; 39BAD5F9E63636F6 CRC64;

Query Match 80.0%; Score 36; DB 2; Length 611;
 Best Local Similarity 62.5%; Pred. No. 33;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVCVODG 8
 DB 187 GRICAOEG 194

RESULT 8
 O9IBG7 PRELIMINARY; PRT; 2327 AA.

AC O9IBG7:
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE KIILIN.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FLOOR PLATE;
 RA Matsui M., Mizuseki K., Nakatani J., Nakatani S., Sasai Y.;
 RT "Xenopus kiilin: A dorsolateral factor containing multiple chordin-type
 repeats secreted from the embryonic midline."
 RT Proc. Natl. Acad. Sci. U.S.A. 97:5291-5296(2000).
 DR EMBL: AB026192; BAA95483.1; -.
 SO SEQUENCE 2327 AA; 255800 MW; 0293109329209983 CRC64;

Query Match 80.0%; Score 36; DB 13; Length 2327;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVCVODG 8
 DB 907 RVCVODG 913

RESULT 9
 O81325 PRELIMINARY; PRT; 475 AA.

AC O81325:
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DE F6N15.2 PROTEIN.
 GN F6N15.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA WASHU;
 RT "The A. thaliana Genome Sequencing Project."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA Ryan E., Edwards J., Pape K.;
 RT "The sequence of A. thaliana F6N15."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA Waterston R.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF069299; AAC19301.1; -
 SQ SEQUENCE 475 AA; 53173 MW; 22A1951B3828A173 CRC64;

Query Match
 Best Local Similarity 77.8%; Score 35; DB 10; Length 475;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRVCVODG 8
 DB 208 RGVCDG 214

RESULT 10.
 ID 09W064 PRELIMINARY; PRT; 549 AA.
 AC 09W064;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GLYCEROL KINASE-LIKE PROTEIN 1.
 GN GK-RS1 OR GKRS1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129SV; TISSUE=TESTES;
 RA Pan Y., Decker W.K., Hug A.H.H.M., Craigen W.J.;
 RT "Retroltransposition of glycerol kinase-related genes from the X
 chromosome to autosomes: Functional and evolutionary aspects."
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF117733; AAD2450.1; -
 DR HSP: P08859; 1GLC.
 DR MCD: MGI:891990; GK-rs1.
 DR INTERPRO: IPR000577; -
 DR INTERPRO: IPR002086; -
 DR PFAM: PF00370; FGGY; 1.
 DR PROSITE: PS00070; ALDEHYDE DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00445; FGGY_KINASES_2; 1.
 DR PROSITE: PS00933; FGGY_KINASES_1; UNKNOWN_1.
 KW kinase.
 SQ SEQUENCE 549 AA; 59871 MW; F807404B79F373F9 CRC64;

Query Match
 Best Local Similarity 77.8%; Score 35; DB 11; Length 549;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
 DB 266 GOLCLDGD 273

RESULT 11

080102
 ID 080102 PRELIMINARY; PRT; 984 AA.
 AC 080102;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE ENDO-N-ACETYLEURAMINIDASE.
 OS bacteriophage #D.
 OC Viruses.
 OX NCBI_TaxID=77920;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Machida Y., Hattori K., Miyake K., Iijima S.;
 RT "Molecular cloning and characterization of a novel bacteriophage-
 associated endo-N-acetylneuraminidase."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB015437; BAA32990.1; -
 DR INTERPRO: IPR001724; -
 DR INTERPRO: IPR002860; -
 DR PFAM: PF02012; BNR; 2.
 DR PRINTS: PR00849; GLHYDRASE58.
 SQ SEQUENCE 984 AA; 108282 MW; 8B09E37C6CD44858 CRC64;

Query Match
 Best Local Similarity 77.8%; Score 35; DB 9; Length 984;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
 DB 643 GSVCKVDG 650

RESULT 12
 ID 09NS98 PRELIMINARY; PRT; 782 AA.
 AC 09NS98;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE SEMAPHORIN SEM2.
 GN SEM2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Saito T., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Miyajima N.,
 RT "Human semaphorin."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB029496; BAA98132.1; -
 SQ SEQUENCE 782 AA; 86700 MW; 85CB424874DF6663 CRC64;

Query Match
 Best Local Similarity 75.6%; Score 34; DB 4; Length 782;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVOD 7
 DB 267 GRVCVND 273

RESULT 13
 ID 069390 PRELIMINARY; PRT; 1048 AA.
 AC 069390;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE DNA POLYMERASE (EC 2.7.7.7) (FRAGMENT).

OS Pseudorabies virus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10345;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-KAPLAN;
 RX MEDLINE=95222727; PubMed=7707503;
 RA Berthomme H., Monahan S.J., Parris D.S., Jacquemont B., Epstein A.L.;
 RT Cloning, sequencing, and functional characterization of the two
 RT subunits of the pseudorabies virus DNA polymerase holoenzyme: evidence
 RT for specificity of interaction.";
 RL J. Virol. 69:2811-2818(1995).
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE - N
 CC PYROPHOSPHATE + DNA(N).
 DR EMBL: L24487; AAA74383.1; -.
 DR INTERPRO: IPR002064; -.
 DR PFAM: PF00136; DNA_pol_B: 1.
 DR PRINTS: PR00106; DNAPOLB.
 DR PROSITE: PS00116; DNA_POLYMERASE_B: 1.
 KW DNA-directed DNA polymerase; DNA replication; DNA-binding.
 FT NON_TER 1048 1048
 SQ SEQUENCE 1048 AA; 115336 MW; B1EB70CFD389276D CRC64;

Query Match 75.3%; Score 34; DB 14; Length 1048;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVCVD 7
 |||||
 DB 475 GRVCVD 481

RESULT 14
 O9KP52 PRELIMINARY; PRT: 185 AA.
 ID O9KP52:
 AC O9KP52: 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE HYPOTHETICAL PROTEIN VC2524.
 GN VC2524.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gmin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayan L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Ragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nieman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae";
 RL Nature 406:477-483(2000).
 DR EMBL: AE004323; AAF95666.1; -.
 DR TIGR: VC2524; -.
 SQ SEQUENCE 185 AA; 20346 MW; F936552E98B04BFA CRC64;

Query Match 73.3%; Score 33; DB 2; Length 185;
 Best Local Similarity 85.7%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RVCVODG 8
 |||||
 DB 136 RVCVADG 142

RESULT 15
 O84353 PRELIMINARY; PRT: 196 AA.
 ID O84353:
 AC O84353: 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE MAF-TYPE PROTEIN.
 GN MAF.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-D/UW-3/CX;
 RX MEDLINE=99000809; PubMed=9784136;
 RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 RT Chlamydia trachomatis.";
 RL Science 282:754-759(1998).
 DR EMBL: AE001308; AAC67944.1; -.
 SQ SEQUENCE 196 AA; 21978 MW; E0B6CFD52F93073 CRC64;

Query Match 73.3%; Score 33; DB 2; Length 196;
 Best Local Similarity 75.0%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRVCVDG 8
 |||||
 DB 150 GRVCVODG 157

Search completed: June 13, 2001, 14:20:22
 Job time: 722 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:16:31 ; Search time 118.55 Seconds
(without alignments)
1.296 Million cell updates/sec

Title: PCT-US01-05825A-2

Perfect score: 45

Sequence: 1 GRVCVQDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCVUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	93.3	400	2 US-08-624-601-8	Sequence 8, Appl1
2	33	73.3	527	4 US-09-311-924-4	Sequence 4, Appl1
3	33	73.3	563	4 US-09-311-924-2	Sequence 2, Appl1
4	33	73.3	745	1 US-08-453-472-5	Sequence 5, Appl1
5	33	73.3	745	1 US-08-038-948-9	Sequence 9, Appl1
6	33	73.3	745	1 US-08-453-952-5	Sequence 5, Appl1
7	33	73.3	745	2 US-08-484-993B-43	Sequence 43, Appl1
8	33	73.3	745	2 US-08-862-903-5	Sequence 5, Appl1
9	33	73.3	745	2 US-08-484-158B-43	Sequence 43, Appl1
10	33	73.3	745	2 US-08-484-596A-43	Sequence 43, Appl1
11	33	73.3	745	2 US-08-480-150A-43	Sequence 43, Appl1
12	33	73.3	745	3 US-08-458-731-43	Sequence 43, Appl1
13	33	73.3	745	3 US-08-149-223A-43	Sequence 43, Appl1
14	32	71.1	326	3 US-09-154-874-9	Sequence 9, Appl1
15	31	68.9	84	2 US-08-465-380-51	Sequence 51, Appl1
16	31	68.9	84	2 US-08-486-397-51	Sequence 51, Appl1
17	31	68.9	84	2 US-08-486-399-51	Sequence 51, Appl1
18	31	68.9	84	2 US-08-461-965-51	Sequence 51, Appl1
19	31	68.9	84	2 US-08-634-641-51	Sequence 51, Appl1
20	31	68.9	84	3 US-09-249-471-51	Sequence 51, Appl1
21	31	68.9	84	3 US-09-249-472-51	Sequence 51, Appl1
22	31	68.9	84	3 US-09-249-451-51	Sequence 51, Appl1
23	31	68.9	84	3 US-08-809-455-51	Sequence 51, Appl1
24	31	68.9	84	3 US-09-249-461-51	Sequence 51, Appl1
25	31	68.9	84	4 US-09-249-448-51	Sequence 51, Appl1
26	31	68.9	162	2 US-08-465-380-64	Sequence 64, Appl1
27	31	68.9	162	2 US-08-486-397-64	Sequence 64, Appl1

ALIGNMENTS

28	31	68.9	162	2	US-08-486-399-64	Sequence 64, Appl1
29	31	68.9	162	2	US-08-461-965-64	Sequence 64, Appl1
30	31	68.9	162	2	US-08-634-641-64	Sequence 64, Appl1
31	31	68.9	162	3	US-09-249-471-64	Sequence 64, Appl1
32	31	68.9	162	3	US-09-249-472-64	Sequence 64, Appl1
33	31	68.9	162	3	US-09-249-451-64	Sequence 64, Appl1
34	31	68.9	162	3	US-08-809-455-64	Sequence 64, Appl1
35	31	68.9	162	3	US-09-249-461-64	Sequence 64, Appl1
36	31	68.9	162	4	US-09-249-448-64	Sequence 64, Appl1
37	31	68.9	327	3	US-09-154-874-8	Sequence 8, Appl1
38	31	68.9	448	2	US-08-811-897A-18	Sequence 18, Appl1
39	31	68.9	448	2	US-08-855-213-18	Sequence 18, Appl1
40	31	68.9	467	2	US-08-811-897A-19	Sequence 19, Appl1
41	31	68.9	467	2	US-08-855-213-19	Sequence 19, Appl1
42	31	68.9	476	2	US-08-811-897A-20	Sequence 20, Appl1
43	31	68.9	495	2	US-08-855-213-20	Sequence 20, Appl1
44	31	68.9	495	2	US-08-811-897A-21	Sequence 21, Appl1
45	31	68.9	495	2	US-08-855-213-21	Sequence 21, Appl1

RESULT 1

US-08-624-601-8
Sequence 8, Application US/08624601
Patent No. 5882653
GENERAL INFORMATION:
APPLICANT: Kaper Dr., James B.
APPLICANT: Levine Dr., Myron B.
TITLE OF INVENTION: Vibrio cholerae O1 (CVD111) and non-O1
TITLE OF INVENTION: Of making and CVD112 and CVD112M) serogroup vaccine strains, methods
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Spencer & Frank
STREET: 1100 New York Ave. N.W. Suite 300 East
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/624,601
FILING DATE: 08-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Scheller Dr., John W.
REGISTRATION NUMBER: 26,031
REFERENCE/DOCKET NUMBER: BAWC20019P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)414-4000
TELEFAX: (202)414-4040
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Vibrio cholerae
STRAIN: El Tor 7946
IMMEDIATE SOURCE:
CLONE: zot
US-08-624-601-8

Query Match 93.3%; Score 42; DB 2; Length 400;
Best Local Similarity 87.5%; Pred. No. 1.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
||:|||||
DB 291 GRVCVODG 298

RESULT 2
US-09-311-924-4
; Sequence 4, Application US/09311924
; Patent No. 6183990
; GENERAL INFORMATION:
; APPLICANT: DUECKER, KLAUS NORBERT
; APPLICANT: CALMELS, THIERRY PAUL GERARD
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30133
; CURRENT APPLICATION NUMBER: US/09/311,924
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: U.K. 9801409.4
; EARLIER FILING DATE: 1998-05-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-311-924-4

Query Match 73.3%; Score 33; DB 4; Length 527;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVOD 7
||||:|
DB 83 GRVCISD 89

RESULT 3
US-09-311-924-2
; Sequence 2, Application US/09311924
; Patent No. 6183990
; GENERAL INFORMATION:
; APPLICANT: DUECKER, KLAUS NORBERT
; APPLICANT: CALMELS, THIERRY PAUL GERARD
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30133
; CURRENT APPLICATION NUMBER: US/09/311,924
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: U.K. 9801409.4
; EARLIER FILING DATE: 1998-05-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-311-924-2

Query Match 73.3%; Score 33; DB 4; Length 563;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVOD 7
||||:|
DB 119 GRVCISD 125

RESULT 4

US-08-453-472-5
; Sequence 5, Application US/08453472
; Patent No. 5626846
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMUNIZATION WITH ZONA PELLUCIDA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & PINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,472
; FILING DATE: 30-May-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,948
; FILING DATE: 26-MAR-1993
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AOTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4032 US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLLE:
; FEATURE:
; NAME/KEY: 2P2
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: human 2P2 protein
US-08-453-472-5

Query Match 73.3%; Score 33; DB 1; Length 745;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
|:|:||||

Db 369 GELCTQDG 376

RESULT 5
US-08-038-948-9
; Sequence 9, Application US/08038948

; Patent No. 5641487

; GENERAL INFORMATION:

; APPLICANT: DEAN, JURRIEN

; TITLE OF INVENTION: CONTRACEPTIVE VACCINE BASED ON

; TITLE OF INVENTION: ALLOIMMUNIZATION WITH ZONA PELLUCIDA POLYPEPTIDES

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN

; STREET: 1100 New York Avenue, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.

; ZIP: 20005-3918

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/038,948

; FILING DATE: 26-MAR-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/930,462

; FILING DATE: 20-AUG-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/364,379

; FILING DATE: 12-JUN-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: SCOTT, Watson T.

; REGISTRATION NUMBER: 26,581

; REFERENCE/DOCKET NUMBER: 99152/E-266-88/2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 861-3000

; TELEFAX: (202) 822-0944

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 745 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-038-948-9

; Query Match 73.3%; Score 33; DB 1; Length 745;

; Best Local Similarity 62.5%; Pred. No. 1.5e+02;

; Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

; Oy 1 GRVCVQDG 8

; Db 369 GELCTQDG 376

; RESULT 6

; US-08-453-952-5

; Sequence 5, Application US/08453952

; Patent No. 5672488

; GENERAL INFORMATION:

; APPLICANT: DEAN, JURRIEN

; TITLE OF INVENTION: CONTRACEPTIVE VACCINE

; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/453,952

FILING DATE: 30-MAY-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/038,948

FILING DATE: 26-MAR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/930,462

FILING DATE: 20-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/364,379

FILING DATE: 12-JUN-1989

ATTORNEY/AGENT INFORMATION:

NAME: DOROTHY R. AUTH

REGISTRATION NUMBER: 36,434

REFERENCE/DOCKET NUMBER: 2026-4032 USA

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 745

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: human

STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

ORGANELLE:

FEATURE:

NAME/KEY: ZP2

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: human ZP2 protein

US-08-453-952-5

; Query Match 73.3%; Score 33; DB 1; Length 745;

; Best Local Similarity 62.5%; Pred. No. 1.5e+02;

; Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

; Oy 1 GRVCVQDG 8

; Db 369 GELCTQDG 376

; RESULT 7

; US-08-484-993B-43

; Sequence 43, Application US/08484993B

; Patent No. 5637497

; GENERAL INFORMATION:

; APPLICANT: Harris Ph.D., Jeffrey D.

; APPLICANT: Hsu, Kuang T.

; APPLICANT: Podolski, Joseph S.

;; TITLE OF INVENTION: Materials and Methods for Immuncontraception
;; NUMBER OF SEQUENCES: 59
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;; STREET: 6300 Sears Tower, 233 South Wacker Drive
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: United States of America
;; ZIP: 60606-6402
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/484,993B
;; FILING DATE: 09-NOV-1993
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/012,990
;; FILING DATE: 29-JAN-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/973,341
;; FILING DATE: 09-NOV-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Clough, David W.
;; REGISTRATION NUMBER: 36,107
;; REFERENCE/DOCKET NUMBER: 31745
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312/474-6653
;; TELEFAX: 312/474-0448
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 43:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 745 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-484-993B-43

Query Match 73.3%; Score 33; DB 2; Length 745;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
|:| | | |
Db 369 GELCTQDG 376

RESULT 8
US-08-862-903-5
; Sequence 5, Application US/08862903
; Patent No. 5916768
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERECT 5.1
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/862,903
;; FILING DATE: 30-May-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/038,948
;; FILING DATE: 26-MAR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/930,462
;; FILING DATE: 20-AUG-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/364,379
;; FILING DATE: 12-JUN-1989
;; ATTORNEY/AGENT INFORMATION:
;; NAME: DOROTHY R. AUTH
;; REGISTRATION NUMBER: 36,434
;; REFERENCE/DOCKET NUMBER: 2026-4032 USA
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 758-4800
;; TELEFAX: (212) 751-6849
;; TELEX: 421792
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 745
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
;; ORIGINAL SOURCE:
;; ORGANISM: human
;; STRAIN:
;; INDIVIDUAL ISOLATE:
;; DEVELOPMENTAL STAGE:
;; HAPLOTYPE:
;; TISSUE TYPE:
;; CELL TYPE:
;; CELL LINE:
;; ORGANELLER:
;; FEATURE:
;; NAME/KEY: ZP2
;; LOCATION:
;; IDENTIFICATION METHOD:
;; OTHER INFORMATION: human ZP2 protein
;; US-08-862-903-5

Query Match 73.3%; Score 33; DB 2; Length 745;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
|:| | | |
Db 369 GELCTQDG 376

RESULT 9
US-08-484-158B-43
; Sequence 43, Application US/08484158B
; Patent No. 5976545
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Pharmaceutical Compositions for
; TITLE OF INVENTION: Immuncontraception
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,1588
FILING DATE: 07-JUNE-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 09-NOV-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-92
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 32794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-1588-43

Query Match
Best Local Similarity 73.3%; Score 33; DB 2; Length 745;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
I : I I I I
Db 369 GELCTODG 376

RESULT 10
US-08-484-596A-43
Sequence 43, Application US/08484596A
Patent No. 5981228
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,596A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 11-NOV-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-596A-43

Query Match
Best Local Similarity 73.3%; Score 33; DB 2; Length 745;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
I : I I I I
Db 369 GELCTODG 376

RESULT 11
US-08-480-150A-43
Sequence 43, Application US/08480150A
Patent No. 5989550
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,150A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/149,223
FILING DATE: 09-NOV-1993
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:

LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-150A-43

Query Match 73.3%; Score 33; DB 3; Length 745;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
| : | | |
Db 369 GELCTODG 376

RESULT 12
US-08-458-731-43
Sequence 43, Application US/08458731
Patent No. 6001599
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,731
FILING DATE: 09-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-458-731-43

Query Match 73.3%; Score 33; DB 3; Length 745;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
| : | | |
Db 369 GELCTODG 376

RESULT 13
US-08-149-223A-43
Sequence 43, Application US/08149223A
Patent No. 6027727
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,223A
FILING DATE: 09-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-149-223A-43

Query Match 73.3%; Score 33; DB 3; Length 745;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVODG 8
| : | | |
Db 369 GELCTODG 376

RESULT 14
US-09-154-874-9
Sequence 9, Application US/09154874
Patent No. 6054636
GENERAL INFORMATION:
APPLICANT: FADER, GARY MICHAEL
TITLE OF INVENTION: ISOFLAVONE BIOSYNTHETIC ENZYMES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA

ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/154,874
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/931,668
FILING DATE: SEPTEMBER 17, 1997
ATTORNEY/AGENT INFORMATION:
NAME: MAJARIAN, WILLIAM R.
REGISTRATION NUMBER: 41,173
REFERENCE/DOCKET NUMBER: BB-1098-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-4926
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 326 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-154-874-9

Query Match 71.1%; Score 32; DB 3; Length 326;
Best Local Similarity 75.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVCVQDC 8
|||||
DB 6 GRVCVTGC 13

RESULT 15
US-08-465-380-51
Sequence 51, Application US/08465380
Patent No. 5863894
GENERAL INFORMATION:
APPLICANT: George P. Vlausk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,380
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 16, 1994

ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/268
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-465-380-51

Query Match 68.9%; Score 31; DB 2; Length 84;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVCVOD 7
|||||
DB 48 GRVCVCD 54

Search completed: June 13, 2001, 14:16:31
Job time: 492 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:14:30 ; Search time 229.28 Seconds
(without alignments)
1.995 Million cell updates/sec

Title: PCT-US01-05825A-3

Perfect score: 41
Sequence: 1 GRVLVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	8	21	Y79107
2	38	92.7	8	21	Y79111
3	37	90.2	118	18	W19878
4	36	87.8	8	21	Y79105
5	34	82.9	8	21	Y84661
6	34	82.9	8	21	Y79115
7	34	82.9	8	21	Y79119
8	34	82.9	20	20	W94489
9	34	82.9	20	21	Y79133
10	34	82.9	25	15	R52775
11	34	82.9	25	15	R52793

12	34	82.9	246	14	R40924	Protein able to b1
13	34	82.9	525	17	R96103	Hepatitis E virus
14	34	82.9	525	17	R96104	Hepatitis E virus
15	34	82.9	540	17	R96101	Hepatitis E virus
16	34	82.9	540	17	R96102	Hepatitis E virus
17	34	82.9	547	13	R26189	Epidemic NANBH vir
18	34	82.9	549	17	R96091	Hepatitis E virus
19	34	82.9	549	17	R96092	Hepatitis E virus
20	34	82.9	549	19	W76367	Hepatitis E virus
21	34	82.9	659	14	R38787	HEV ORF2 protein.
22	34	82.9	659	14	R39308	Mexico strain HEV
23	34	82.9	659	18	W35827	Hepatitis E virus
24	34	82.9	659	20	W93387	Human HEV ORF 2 pr
25	34	82.9	660	12	R14619	Protein encoded by
26	34	82.9	660	14	R38785	HEV ORF2 protein.
27	34	82.9	660	14	R39306	Burma strain HEV O
28	34	82.9	660	15	R51265	HEV strain protein
29	34	82.9	660	16	R70323	Hepatitis E virus
30	34	82.9	660	17	R91814	Hepatitis E virus
31	34	82.9	660	17	R96089	Hepatitis E virus
32	34	82.9	660	17	R96090	Hepatitis E virus
33	34	82.9	660	18	W35826	Hepatitis E virus
34	34	82.9	660	19	W81520	Hepatitis E virus
35	34	82.9	660	19	W80197	Protein encoded by
36	34	82.9	660	19	W76369	Hepatitis E virus
37	34	82.9	660	19	W71210	Protein encoded by
38	34	82.9	660	20	Y31385	HEV-ORF2 prote
39	34	82.9	660	20	Y31382	HEV-ORF2 prote
40	34	82.9	660	20	W93386	Human HEV ORF 2 pr
41	34	82.9	660	20	W93388	Human HEV ORF 2 pr
42	34	82.9	660	20	W93389	Human HEV ORF 2 pr
43	34	82.9	660	20	W93390	Human HEV ORF 2 pr
44	34	82.9	660	20	W93391	Human HEV ORF 2 pr
45	34	82.9	660	20	W93392	Human HEV ORF 2 pr

ALIGNMENTS

RESULT 1	
Y79107	Y79107 standard; Peptide; 8 AA.
XX	XX
AC	Y79107;
XX	XX
DT	05-JUN-2000 (first entry)
XX	XX
DE	Peptide antagonist of zonulin.
XX	XX
KW	zonulin; antagonist; zonula occludens toxin receptor;
KW	blood-brain barrier; antiinflammatory; cerebroprotective;
KW	neuroprotective; dermatological; antifungal; antiviral;
KW	antibacterial; cytostatic; anti-HIV; vulnerary; antiallergic;
KW	hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW	gastrointestinal inflammation; therapy.
XX	XX
OS	Synthetic.
XX	XX
PN	W0200007609-A1.
XX	XX
PD	17-FEB-2000.
XX	XX
PE	28-JUL-1999; 99WO-US16683.
XX	XX
PR	03-AUG-1998; 98US-0127815.
XX	XX
PA	(UYMA-) UNIV MARYLAND BALTIMORE.
XX	XX
PI	Fasano A;
XX	XX
DR	WPI; 2000-205565/18.
XX	XX
PT	New peptide antagonist of zonulin useful as antiinflammatory agent for

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
XX
XX
PS Claim 1; Page 41; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.
CC
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 41; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLVQPG 8
|||
1 grvlvpg 8

RESULT 2
Y79111
ID Y79111 standard; Peptide; 8 AA.
XX
AC Y79111;
XX

DT 05-JUN-2000 (first entry)
XX

DE Peptide antagonist of zonulin.
XX

KW Zonulin; antagonist; zonula occludens toxin receptor;
KW blood-brain barrier; antiinflammatory; cerebroprotective;
KW neuroprotective; dermatological; antitumor; antiviral;
KW antibacterial; cytotoxic; anti-HIV; vulnerrary; anti-allergic;
KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW gastrointestinal inflammation; therapy.
XX

OS Synthetic.
XX

PN WO200007609-A1.
XX

PD 17-FEB-2000.
XX

PF 28-JUL-1999; 99WO-US16683.
XX

PR 03-AUG-1998; 98US-0127815.
XX

PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX

PI Fasano A;
XX
DR WPI; 2000-205565/18.
XX

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
XX
XX
PS Claim 1; Page 43; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor, yet do
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.
CC
XX
SQ Sequence 8 AA;

Query Match 92.7%; Score 38; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. No. 3.2e+05;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLVQPG 8
|||
1 grvlvpg 8

RESULT 3
W19878
ID W19878 standard; Protein; 118 AA.
XX
AC W19878;
XX

DT 07-DEC-1997 (first entry)
XX

DE CEA-specific antibody CEA3 VH sequence.
XX

KW Carcinoembryonic antigen; CEA; human; antibody; scfv;
KW tumour marker; lung cancer; breast cancer; colon cancer;
KW adenocarcinoma; diagnosis.
XX

OS Homo sapiens.
XX

EH Key
FT Location/Qualifiers
FT 31..35

FT /label= CDR1
FT /note= "complementarity determining region 1"
FT 50..66

FT /label= CDR2
FT /note= "complementarity determining region 2"
FT 99..107

FT Region

```

FT      /label= CDR3
FT      /note= "Complementarity determining region 3"
PN      WO9720932-A1..
PD      12-JUN-1997.
XX
XX
XX      09-DEC-1996;    96WO-GB03043.
XX
XX      11-OCT-1996;    96GB-0021295.
PR      07-DEC-1995;    95GB-0025004.
PR      23-MAY-1996;     96GB-0010824.
XX
XX      (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
XX      Allen DJ, McCaferly JG, Osbourn JK;
PI
XX      WPI: 1997-319779/29.
DR      N-PSDB: T72128.
XX
XX      Specific binding members for human carcinoembryonic antigen - bind
PT      to the A3-B3 extracellular domain of hCEA and are substantially
PT      non-cross-reactive with human liver cells; used for diagnosing
PT      cancer
XX
XX      Claim 7: Fig 1a; 128pp; English.
XX
XX      This polypeptide sequence comprises the heavy chain variable
CC      region (VH) of human carcinoembryonic antigen (hCEA)-specific
CC      antibody CEA3. VH (T72126-32) and VL (T72133-35) gene sequences
CC      were obtained for anti-hCEA antibodies CEA1-CEA7 (see W19876-85).
CC      A claimed specific binding member (A) comprises an hCEA specific
CC      antibody antigen binding domain that has a dissociation constant
CC      for hCEA of less than 1 x 10-8 M. is non-cross-reactive with human
CC      liver cells, and preferentially binds to the A3-B3 extracellular
CC      domain of hCEA and/or to cell-associated hCEA over hCEA over
CC      soluble hCEA. Preferred (A) include pairings of VH and VL
CC      sequences from CEA1-7, or their CDR sequences, as well as CEA6
CC      VH and VL variants. (A) is used to detect cells expressing hCEA,
CC      in vivo or in vitro, especially tumor cells for diagnosing cancer,
CC      e.g. adenocarcinoma of the colon, lung or breast.
XX
XX      Sequence 118 AA:
SQ
XX
XX      Query Match          90.2%; Score 37; DB 18; Length 118;
XX      Best Local Similarity 87.5%; Pred. No. 6.8;
XX      Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0.
OY      1 GRVLQPG 8
XX      |||||
DB      8 grrlvpgp 15

RESULT 4
Y79105
ID      Y79105 standard; Peptide; 8 AA.
XX
XX      Y79105;
XX
XX      05-JUN-2000 (first entry)
XX
XX      Peptide antagonist of zonulin.
XX
XX      Zonulin; antagonist; zonula occludens toxin receptor;
KM      blood-brain barrier; antiinflammatory; cerebroprotective;
KM      neuroprotective; dermatological; antiulcer; antilivral;
KM      antibacterial; cytostatic; anti-HIV; vulnerary; antiallergic;
KM      hypotensive; immunosuppressive; antiparasitic; vasotropic;
KM      gastrointestinal inflammation; therapy.
XX
XX      Synthetic.
XX

```

PN WO200007609-A1.
XX
XX PD 17-FEB-2000.
XX
PF 28-JUL-1999; 99WO-US16683.
XX
XX 03-AUG-1998; 98US-O127815.
XX
PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PI Fasano A;
DR WPI; 2000-205565/18.
XX
XX New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
XX
XX PS Claim 1; Page 41; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
CC (Z), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.

XX Sequence 8 AA:
SQ

Query Match 87.8%; Score 36; DB 21; Length 8;
Best Local Similarity 87.5%; Pred No. 3.2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLQVP 8
||| ||||
DB 1 grvcvqpg 8

RESULT 5
ID Y84661 Y84661 standard; Protein; 8 AA.
AC Y84661;
XX
XX 25-JUL-2000 (first entry)
DE Peptide antagonist FZI/O of zonula occludens toxin (zot) polypeptide.

XX Human; Zot: zonula occludens toxin; zonulin; antigen presenting cell;
KW APC; lymphocyte proliferation; antigen; auto-immune disorder;
KW immune-related disorder; immune system rejection; multiple sclerosis;
KW organ transplantation; inflammatory disease; allergic disease;

KW Rheumatoid arthritis; insulin dependent diabetes mellitus;
 KW celiac disease; Sjogren's syndrome; systemic lupus erythematosus;
 KW auto-immune thyroiditis; idiopathic thrombocytopenic purpura;
 KW hemolytic anemia; Grave's disease; autoimmune orchitis;
 KW pernicious anemia; vasculitis; Addison disease; autoimmune orchitis;
 KW myasthenia gravis; polynuritis; autoimmune coagulopathy; polymyositis;
 KW Dermatomyositis; scleroderma; asthma; psoriasis; eczematous dermatitis;
 KW Kaposi's sarcoma; inflammatory bowel disease; proliferative disorder;
 PCR primer; ss.

Synthetic.

WO200015252-A1.

23-MAR-2000.

09-SEP-1999; 99WO-US18842.

14-SEP-1998; 98US-0100266.

(UYMA-) UNIV MARYLAND BALTIMORE.

Fasano A, Stein MB, Lu R, Tanner MK;

WPI; 2000-271257/23.

Suppression of antigen presenting cell mediated lymphocyte proliferation, by administering a Zot-related immunoregulator useful for treating immune-related disorders, immune system rejection subsequent to tissue or organ transplantation

Example 7; Page 59; 95pp; English.

The present sequence represents a peptide antagonist of zonula occludens toxin (Zot) polypeptide. The specification describes a method of suppressing antigen presenting cell (APC)-mediated lymphocyte proliferation in a mammalian host pre-exposed to a particular antigen. The method comprises administering to the host an effective amount of a Zot-related immunoregulator selected from Zot (zonula occludens toxin) or zonulin, the amount effective to down-regulate the activity of the APC. The method can be used to down-regulate APC-mediated lymphocyte proliferation in mammalian hosts suffering from auto-immune or immune-related disorders, immune system rejection subsequent to tissue or organ transplantation, or inflammatory or allergic diseases. The autoimmune or immune related disorders include multiple sclerosis, rheumatoid arthritis, insulin dependent diabetes mellitus, celiac disease, Sjogren's syndrome, systemic lupus erythematosus, autoimmune thyroiditis, idiopathic thrombocytopenic purpura, hemolytic anemia, Grave's disease, Addison disease, autoimmune orchitis, pernicious anemia, vasculitis, autoimmune coagulopathy, myasthenia gravis, dermatomyositis, pemphigus, rheumatic carditis, polymyositis, Dermatomyositis, and scleroderma. The inflammatory or allergic disease Kaposi's sarcoma, multiple sclerosis, inflammatory bowel disease, CC proliferative disorders of smooth muscle cells, and inflammatory CC conditions associated with mycotic, viral, parasitic, or bacterial infections.

SO Sequence 8 AA;

Query Match 82.9%; Score 34; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVLVPG 8
 I I I I I I I I
 Db 1 gylvlvpq 8

RESULT 6
 Y79115
 ID Y79115 standard; Peptide; 8 AA.

XX
 AC Y79115;

05-JUN-2000 (first entry)

Peptide antagonist of zonulin.

KW Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; anti-inflammatory; cerebroprotective;
 KW neuroprotective; dermatological; anticancer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; antiviral; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.

Synthetic.

WO200007609-A1.

17-FEB-2000.

28-JUL-1999; 99WO-US16683.

03-AUG-1998; 98US-0127815.

(UYMA-) UNIV MARYLAND BALTIMORE.

Fasano A;

WPI; 2000-205565/18.

New peptide antagonist of zonulin useful as anti-inflammatory agent for treating cerebral ischemia, stroke, cerebral edema, gastritis, shigellosis, viral gastroenteritis, meningitis, encephalomyelitis

Claim 1; Page 44; 69pp; English.

This present sequence is that of a peptide antagonist of zonulin (Z), one of 25 such peptides (see Y79105-29) of the invention, which bind to a zonula occludens toxin (ZOT) receptor, yet do not physiologically modulate the opening of mammalian tight junctions (TJ). The peptide antagonists are based on a common motif of ZOT binding and human zonulins, which is believed to be critical for receptor binding. They can be prepared by chemical synthesis or by use of recombinant DNA techniques. The peptide antagonists are used as anti-inflammatory agents in the treatment of gastrointestinal inflammation, where they bind to the ZOT receptor in the intestine and yet does not physiologically modulate the opening of TJ in the intestine. Gastrointestinal inflammation conditions give rise to treating intestinal permeability and the peptide is useful for caused by infection, e.g. Clostridium difficile infection, enterocolitis, shigellosis, viral gastroenteritis, parasite infection, bacterial overgrowth, whipple's disease, diseases with mucosal erosion or ulcerations, e.g. gastritis, gastric cancer, collagenous colitis, inflammatory bowel disease, diseases marked by lymphatic obstruction, e.g. congenital intestinal lymphangectasia, sarcooidosis lymphoma, mesenteric tuberculosis, and after surgical correction of congenital heart disease with Fontan's operation, mucosal diseases without ulceration, e.g. Menetrier's disease, celiac disease, eosinophilic gastroenteritis, and immune diseases, e.g. systemic lupus erythematosus or food allergies, primarily to milk.

SO Sequence 8 AA;

Query Match 82.9%; Score 34; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVLVPG 8
 I I I I I I I I
 Db 1 gylvlvpq 8

RESULT 7

Y79119 standard; Peptide; 8 AA.

Y79119;

05-JUN-2000 (first entry)

Peptide antagonist of zonulin.

Zonulin: antagonist; zonula occludens toxin receptor; human; blood-brain barrier; antiinflammatory; cerebroprotective; neuroprotective; dermatological; antiulcer; antiviral; antibacterial; cytostatic; anti-HIV; vulnery; antiallergic; hypotensive; immunosuppressive; antiparasitic; vasotropic; gastrointestinal inflammation; therapy.

Homo sapiens.

WO200007609-A1.

17-FEB-2000.

28-JUL-1999; 99WO-US16683.

03-AUG-1998; 98US-0127815.

(UYMA-) UNIV MARYLAND BALTIMORE.

Faasano A;

WPI: 2000-205565/18.

New peptide antagonist of zonulin useful as antiinflammatory agent for treating cerebral ischemia, stroke, cerebral edema, gastritis, shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -

Claim 1: Page 45; 69pp; English.

This present sequence is that of a peptide antagonist of zonulin, corresponding to residues 8-15 of human foetal intestinal zonulin. It is one of 25 peptide antagonists (see Y79105-29) of the invention that bind to a zonula occludens toxin (ZOT) receptor, yet do not physiologically modulate the opening of mammalian tight junctions (TJ). The peptide antagonists are based on a common motif of ZOT and human zonulins, which is believed to be critical for receptor binding. They can be prepared by chemical synthesis or by use of recombinant DNA techniques. The peptide antagonists are used as an antiinflammatory agents in the treatment of gastrointestinal inflammation, where they bind to the ZOT receptor in the intestine and yet does not physiologically modulate the opening of TJ in the intestine. Gastrointestinal inflammation conditions give rise to increased intestinal permeability and the peptide is useful for treating intestinal conditions that cause protein losing enteropathy, caused by infection, e.g. Clostridium difficile infection, enterocolitis, shigellosis, viral gastroenteritis, parasite infestation, bacterial overgrowth, whipple's disease, diseases with mucosal erosion or ulcerations, e.g. gastritis, gastric cancer, collagenous colitis, inflammatory bowel disease, diseases marked by lymphatic obstruction, e.g. congenital intestinal lymphangiectasia, bartolodosis lymphoma, mesenteric tuberculosis, and after surgical correction of congenital heart disease with Fontan's operation, mucosal diseases without ulceration, e.g. Menetrier's disease, coeliac disease, eosinophilic gastroenteritis, and immune diseases, e.g. systemic lupus erythematosus or food allergies, primarily to milk.

Sequence 8 AA;

Query Match 82.9%; Score 34; DB 21; Length 8;

Best Local Similarity 87.5%; Pred. No. 3.2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GRVLVPG 8

Db 1 ggvlvpgp 8

RESULT 8

W94489 standard; peptide; 20 AA.

W94489;

21-APR-1999 (first entry)

Human foetal intestine zonulin N-terminal peptide.

Zonulin; mammalian tight junction; zonula occludens toxin; ZOT;

Vibrio cholerae; vaccine; cholera toxin; polyclonal antibody; intestinal mucosa; nasal mucosa; blood brain barrier.

Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 16 /note= "unspecified"

WO9852415-A1.

26-NOV-1998.

28-APR-1998; 98WO-US07636.

21-MAY-1997; 97US-0859931.

(UYMA-) UNIV MARYLAND BALTIMORE.

Faasano A;

WPI: 1999-070123/06.

New purified zonulin - which is capable of reversibly opening mammalian tight junctions, used for enhancing the delivery of agents across intestinal and nasal mucosa and blood brain barrier

Claim 2: Page 45; 64pp; English.

The present invention describes pure zonulin which has an apparent molecular weight of 47 kD, as determined by SDS-PAGE, which is recognised by both anti-tau polyclonal antibody and by anti-zonula occludens toxin (ZOT) polyclonal antibody, and is capable of reversibly opening mammalian tight junctions. Zonulin proteins function as physiological modulators of mammalian tight junctions. They can be used for enhancing the absorption of therapeutic agents across tight junctions of intestinal and nasal mucosa and across tight junctions of the blood brain barrier. Zonulin can be used with agents such as drugs, e.g. lidocaine, adenosine, dobutamine, dopamine, epinephrine, norepinephrine, phenolamine, doxapram, alfentanil, dezocin, nalbuphine, buprenorphine, naloxone, ketorolac, midazolam, propofol, metacurine, vincuristine, succinylcholine, cytarabine, mitomycin doxorubicin, vincristine, vinblastine, methicillin, mezlocillin, piperacillin, cefoxitin, cefenicol, cefmetazole and aztreonam, a hormone e.g. testosterone, nandrolone, menotropins, insulin, urofollitropin, interferon-alpha, interferon-beta, interferon-gamma, interleukin-1 (IL-1), IL-2, IL-4, IL-8, polyvalent IgG, specific IgG, IgA, or IgM. The proteins can also be used for the production of antibodies which can be used to assay for zonulin in body tissue or fluids, or in affinity-purification of zonulin. The present sequence represents an N-terminal peptide of zonulin.

Sequence 20 AA;

Query Match 82.9%; Score 34; DB 20; Length 20;
Best Local Similarity 87.5%; Pred. No. 4.4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLYOPG 8
| | | | |
Db 8 ggvlyvpg 15

RESULT 9
Y79133
ID Y79133 standard; Peptide: 20 AA.
XX
AC Y79133;
XX
DT 05-JUN-2000 (first entry)
XX
DE Human foetal intestine zonulin N-terminal sequence.
XX
KW Zonulin: antagonist; zonula occludens toxin receptor;
KM human; blood-brain barrier; antiinflammatory;
KN gastrointestinal inflammation; therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 16 /note= "unidentified residue"
FT
XX
FN W0200007609-A1.
XX
PD 17-FEB-2000.
XX
PF 28-JUL-1999; 99MO-US16683.
XX
PR 03-AUG-1998; 98US-0127815.
XX
PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PI Fasano A;
XX
DR WPI; 2000-205565/18.
XX
PT New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
XX
PS Example 3; Fig 6; 69pp; English.
XX
CC The present sequence is that of the N-terminal region of foetal
CC human intestinal zonulin. The N-terminal sequences of human adult
CC and foetal zonulins (see Y79130-36) were compared with Vibrio cholerae
CC zonula occludens toxin (ZOT) to identify a common motif thought
CC to be involved in receptor binding. Peptide antagonists (see
CC Y79105-29) based on this motif are useful as antiinflammatory
CC agents for treatment of gastrointestinal inflammation, and for
CC treatment of conditions associated with breakdown of the blood-brain
CC barrier.
XX
SQ Sequence 20 AA;

Query Match 82.9%; Score 34; DB 21; Length 20;
Best Local Similarity 87.5%; Pred. No. 4.4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLYOPG 8
| | | | |
Db 8 ggvlyvpg 15

RESULT 10

R52775
ID R52775 standard; Protein: 25 AA.
XX
AC R52775;
XX
DT 24-JAN-1995 (first entry)
XX
DE Murine Bre-3 immunoglobulin heavy chain variable domain N-terminus.
XX
KW Immunoglobulin variable domain; primer; polymerase chain reaction;
KM chimeric antibody; human milk fat globule; Bre-3 VL-chain.
XX
OS Mus musculus.
XX
PN W09411508-A.
XX
PD 26-MAY-1994.
XX
PF 15-NOV-1993; 93MO-US11316.
XX
PR 13-NOV-1992; 92US-0977706.
PR 13-NOV-1992; 92US-0977707.
PR 28-SEP-1993; 93US-0128015.
XX
PA (CANC-) CANCER RES FUND CONTRA COSTA.
XX
DR WPI; 1994-183509/22.
XX
PT Chimeric human-murine polypeptide(s) specific for human mammary
PT fat globule antigen - for imaging, diagnosing and treating
PT neoplasia, with less undesirable immunogenic response
XX
PS Example 11; Page 32; 54pp; English.
XX
CC Primers J02, J03, J04, J014 and VH1BACK (062740-062744) were all
CC used to prepare cDNAs that encode the Bre-3 mouse Ig variable
CC domains. The amplified V-regions lacked constant regions so as to
CC produce less immunogenic polypeptides. A hybrid polypeptide was
CC prepared using human constant regions with the murine V regions.
CC The chimeric polypeptide retained the binding affinity of Bre-3
CC for human milk fat globule. The amino acid sequences deduced from
CC sequences which had been determined directly (i.e. R52774 and
CC R52775). The general agreement between the predicted and the
CC determined amino acid sequences that the cloned cDNAs encode Bre-3.
XX
SQ Sequence 25 AA;

Query Match 82.9%; Score 34; DB 15; Length 25;
Best Local Similarity 87.5%; Pred. No. 5.4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLYOPG 8
| | | | |
Db 8 ggvlyvpg 15

RESULT 11
R52793
ID R52793 standard; Protein: 25 AA.
XX
AC R52793;
XX
DT 24-JAN-1995 (first entry)
XX
DE Murine Bre-3 immunoglobulin heavy chain variable domain N-terminus.
XX
KW Immunoglobulin variable domain; primer; polymerase chain reaction;
KM chimeric antibody; human milk fat globule; Bre-3 VL-chain.
XX
OS Mus musculus.
XX

PN W09411509-A.
 XX
 XX 26-MAY-1994.
 XX
 XX
 PF 16-NOV-1993; 93WO-US11445.
 XX
 PR 16-NOV-1992; 92US-0977696.
 PR 30-SEP-1993; 93US-0129930.
 PR 08-OCT-1993; 93US-0134346.
 XX
 PA (CANC-) CANCER RES FUND CONTRA COSTA.
 XX
 DR WPI; 1994-183510/22.
 XX
 XX New analogue peptide(s) comprising antibody variable regions -
 PT used to develop prods. for use in the detection, diagnosis,
 PT therapy and prevention of neoplasms
 XX
 PS Example 12; Page 53; 54pp; English.
 XX
 CC Primers J02, J03, J04, J014 and VH1BACK (Q62765-Q62769) were all
 CC used to prepare cDNAs that encode the Bre-3 mouse Ig variable
 CC domains. The amplified V-regions lacked constant regions so as to
 CC produce less immunogenic polypeptides. A hybrid polypeptide was
 CC prepared using human constant regions with the murine V regions.
 CC The chimeric polypeptide retained the binding affinity of Bre-3
 CC for human milk fat globule. The amino acid sequences deduced from
 CC the amplified VL and VH regions were compared to N-terminal
 CC sequences which had been determined directly (i.e. R52792 and
 CC R52793). The general agreement between the predicted and the
 CC determined amino acid sequences that the cloned cDNAs encode Bre-3.
 XX
 SO Sequence 25 AA;

Query Match 82.9%; Score 34; DB 15; Length 25;
 Best Local Similarity 87.5%; Pred. No. 5.4;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVQPG 8
 I I I I I I I
 DB 8 gylvlvpqg 15

RESULT 12
 ID R40924 standard; Protein; 246 AA.
 XX
 AC R40924;
 XX
 DT 17-FEB-1994 (first entry)
 XX
 DE Protein able to bind to HIV-1 tat protein.
 XX
 KW Antibodies; antigen binding proteins; library; HIV;
 KM Human Immunodeficiency Virus.
 XX
 OS Synthetic.
 XX
 PN EP557897-A.
 XX
 PD 01-SEP-1993;
 XX
 PF 19-FEB-1993; 93EP-0102609.
 XX
 PR 28-FEB-1992; 92US-0843125.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Dillon PJ, Rosen CA;
 XX
 DR WPI; 1993-274375/35.
 N-PSDB; Q48605.

XX Antigen-binding proteins and corresp. synthetic genes - are
 PT constructed synthetically and used to make antibody library, in
 PT disease diagnosis, etc.
 XX
 XX
 PS Claim 13; Page 22; 40pp; English.
 XX
 CC The synthetic gene encodes a protein corresponding to an antibody
 CC capable of binding to a specific antigen, in this case the HIV-1 tat
 CC protein. Many synthetic genes are synthesised, each containing a
 CC predetermined nucleotide region encoding the framework regions of
 CC the heavy and light chains of antibody and underdetermined nucleotide
 CC regions which are random sequences. The genes are then used in the
 CC construction of vectors which are subsequently used to transform
 CC microbes. The proteins thus produced are screened for binding
 CC activity to the specific antigen.
 XX
 SO Sequence 246 AA;

Query Match 82.9%; Score 34; DB 14; Length 246;
 Best Local Similarity 87.5%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVQPG 8
 I I I I I I I
 DB 8 gylvlvpqg 15

RESULT 13
 ID R96103 standard; Protein; 525 AA.
 XX
 AC R96103;
 XX
 DT 06-AUG-1996 (first entry)
 XX
 DE Hepatitis E virus (Burma strain) recombinant 62k antigen.
 XX
 KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
 KM diagnosis; antigen; Spodoptera frugiperda; Sf9; insect;
 KW baculovirus; capsid.
 XX
 OS Hepatitis E virus Burma strain.
 XX
 PN W09612807-A2.
 XX
 PD 02-MAY-1996.
 XX
 PF 23-OCT-1995; 95WO-US13703.
 XX
 PR 13-OCT-1995; 95US-0542634.
 PR 24-OCT-1994; 94US-0327952.
 XX
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 XX
 PI Fuerst TR, McAtee CP, Yarbough PO, Zhang Y;
 PI WPI; 1996-230608/23.
 DR Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
 PT diagnostic reagents for determining HEV infection and in vaccines
 PT
 PS Claim 7; Page 95-97; 125pp; English.
 XX
 CC DNA (T27109) coding for the 62k antigen (R96091) of hepatitis E virus
 CC (HEV) Burma strain capsid protein was cloned into baculovirus
 CC expression vector pBlueBac11 and recombinant 62k was expressed
 CC in Sf9 insect cells. High levels of expression were obtd. and
 CC the recombinant 62k was obtd. in over 95% purity. However, C-terminal
 CC processing resulted in the deletion of 9 or 23 amino acids from 62k,
 CC giving 2 related polypeptide species (R96101 and R96103). Similar
 CC results were obtd. with HEV Mexico 62k antigen (see also R96102 and

CC R96104). Recombinant 62K represents an improved antigen, in
 CC comparison to bacterial expressed proteins, for use in HEV
 CC diagnostic assays, and also has excellent immunogenic properties.
 XX
 SQ Sequence 525 AA;

Query Match 82.9%; Score 34; DB 17; Length 525;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RVLVOPG 8
 Db 122 rllvpg 128

RESULT 14
 R96104
 ID R96104 standard; Protein: 525 AA.
 XX
 AC R96104;
 XX
 DT 06-AUG-1996 (first entry)
 XX
 DE Hepatitis E virus (Mexico strain) recombinant 62K antigen.
 XX
 KM HEV: enterically-transmitted non-A/non-B hepatitis virus; vaccine;
 KM diagnosis; antigen; Spodoptera frugiperda; Sf9; Insect;
 KM baculovirus; capsid.
 XX
 OS Hepatitis E virus Mexico strain.
 XX
 PN WO9612807-A2.
 XX
 PD 02-MAY-1996.
 XX
 PF 23-OCT-1995; 95WO-US13703.
 XX
 PR 13-OCT-1995; 95US-0542634.
 XX
 PR 24-OCT-1994; 94US-0327952.
 XX
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 XX
 PI Fuerst TR, McAtee CP, Yarbough PO, Zhang Y;
 XX
 DR WPI; 1996-230608/23.
 XX
 PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
 XX diagnostic reagents for determining HEV infection and in vaccines
 XX
 PS Claim 8; Page 97-98; 125pp; English.
 XX
 CC DNA (T72110) coding for the 62K antigen (R96092) of hepatitis E virus
 CC (HEV) Mexico strain capsid protein was cloned into baculovirus
 CC expression vector pBluscript and recombinant 62K was expressed
 CC in Sf9 insect cells. High levels of expression were obtained, and
 CC the recombinant 62K was obtained. In over 95% purity. However, C-terminal
 CC processing resulted in the deletion of 9 or 23 amino acids from 62K,
 CC giving 2 related polypeptide species (R96102 and R96104). Similar
 CC results were obtained with HEV Burma 62K antigen (see also R96101 and
 CC R96103). Recombinant 62K represents an improved antigen, in
 CC comparison to bacterial expressed proteins, for use in HEV
 CC diagnostic assays, and also has excellent immunogenic properties.
 CC
 SQ Sequence 525 AA;

Query Match 82.9%; Score 34; DB 17; Length 525;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RVLVOPG 8
 Db 122 rllvpg 128

Db 122 rllvpg 128

RESULT 15
 R96101
 ID R96101 standard; Protein: 540 AA.
 XX
 AC R96101;
 XX
 DT 06-AUG-1996 (first entry)
 XX
 DE Hepatitis E virus (Burma strain) recombinant 62K antigen.
 XX
 KM HEV: enterically-transmitted non-A/non-B hepatitis virus; vaccine;
 KM diagnosis; antigen; Spodoptera frugiperda; Sf9; Insect;
 KM baculovirus; capsid.
 XX
 OS Hepatitis E virus Burma strain.
 XX
 PN WO9612807-A2.
 XX
 PD 02-MAY-1996.
 XX
 PF 23-OCT-1995; 95WO-US13703.
 XX
 PR 13-OCT-1995; 95US-0542634.
 XX
 PR 24-OCT-1994; 94US-0327952.
 XX
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 XX
 PI Fuerst TR, McAtee CP, Yarbough PO, Zhang Y;
 XX
 DR WPI; 1996-230608/23.
 XX
 PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
 XX diagnostic reagents for determining HEV infection and in vaccines
 XX
 PS Claim 5; Page 92-93; 125pp; English.
 XX
 CC DNA (T72109) coding for the 62K antigen (R96091) of hepatitis E virus
 CC (HEV) Burma strain capsid protein was cloned into baculovirus
 CC expression vector pBluscript and recombinant 62K was expressed
 CC in Sf9 insect cells. High levels of expression were obtained, and
 CC the recombinant 62K was obtained. In over 95% purity. However, C-terminal
 CC processing resulted in the deletion of 9 or 23 amino acids from 62K,
 CC giving 2 related polypeptide species (R96101 and R96103). Similar
 CC results were obtained with HEV Mexico 62K antigen (see also R96102 and
 CC R96104). Recombinant 62K represents an improved antigen, in
 CC comparison to bacterial expressed proteins, for use in HEV
 CC diagnostic assays, and also has excellent immunogenic properties.
 CC
 SQ Sequence 540 AA;

Query Match 82.9%; Score 34; DB 17; Length 540;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RVLVOPG 8
 Db 122 rllvpg 128

Search completed: June 13, 2001, 14:14:31
 Job time: 372 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:10:37 ; Search time 130.61 Seconds
(without alignments)
4.209 Million cell updates/sec

Title: PCT-US01-05825A-3

Perfect score: 41

Sequence: 1 GRVLPQG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR.67:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	82.9	253	2	B82748
2	34	82.9	396	2	D26956
3	34	82.9	403	2	E71300
4	34	82.9	659	1	B44212
5	34	82.9	660	1	VHMH2
6	34	82.9	1400	2	B70963
7	34	82.9	1440	2	B74872
8	34	80.5	267	2	B82694
9	33	80.5	321	2	T12497
10	33	80.5	322	2	G72643
11	33	80.5	354	2	E82850
12	33	80.5	394	2	S62726
13	33	80.5	405	2	B70300
14	33	80.5	405	2	G70465
15	32	78.0	132	1	G69256
16	32	78.0	354	2	D71539
17	32	78.0	354	2	B81694
18	32	78.0	360	2	F72094
19	32	78.0	569	2	M43317
20	32	78.0	614	2	B71551
21	32	78.0	614	2	H81703
22	32	78.0	655	2	A83395
23	32	78.0	926	2	T48391
24	31	75.6	94	2	S42185
25	31	75.6	98	2	S26928
26	31	75.6	110	2	B72730
27	31	75.6	116	2	I84704
28	31	75.6	120	2	S44111
29	31	75.6	197	2	S22010

ALIGNMENTS

30	31	75.6	238	2	S76936	hypothetical prote
31	31	75.6	308	2	B75292	glucokinase - dein
32	31	75.6	358	2	T36415	probable iron-side
33	31	75.6	375	2	JC7287	G-protein coupled
34	31	75.6	389	2	S31123	hypothetical prote
35	31	75.6	397	2	I40216	translation elonga
36	31	75.6	410	2	E75208	probable valine--p
37	31	75.6	427	2	D83347	probable aminotran
38	31	75.6	439	2	S75545	hypothetical prote
39	31	75.6	447	2	C83683	phosphoglucosamine
40	31	75.6	474	2	A36240	pyrimidine synthe
41	31	75.6	494	2	S30187	mannose-6-phosphat
42	31	75.6	504	2	T10698	legumin-like prote
43	31	75.6	565	2	T10696	legumin-like prote
44	31	75.6	588	2	T35549	hypothetical prote
45	31	75.6	619	2	D81556	conserved hypothet

RESULT 1
B82748
ubiquinol cytochrome C oxidoreductase, cytochrome C1 subunit XF0910 [imported] - Xyle
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82748
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: AB2515; PMID:20365717
A:Note: For a complete list of authors see reference number A59328 below
A:Accession: B82748
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-253 <SIM>
A:Cross-references: GB:AE003930; GB:AE003849; NID:g9105819; PIDN:AAF83720.1; GSPDB:GN
A:Experimental source: strain 945c
R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
R:Riones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincanli, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
Chado, M.A.; Madela, A.M.B.N.; Madela, H.M.F.; Martino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0910

Query Match 82.9%; Score 34; DB 2; Length 253;
Best Local Similarity 87.5%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVLPQG 8
DB 187 GLVLPQG 194

RESULT 2
D26956
translation elongation factor EF-Tu - Micrococcus luteus
C:Species: Micrococcus luteus, Micrococcus lysodeikticus
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 21-Jul-2000
C:Accession: D26956
R:Ohama, T.; Yamao, F.; Muto, A.; Osawa, S.
J. Bacteriol. 169, 4770-4777, 1987

A>Title: Organization and codon usage of the streptomycin operon in *Micrococcus luteus*,
A:Reference number: A91844; MUID:88007427
A:Accession: D26956

A:Molecule type: DNA
A:Residues: 1-396 <CONA>

A:Cross-references: GB:M17788; NID:g149850; PIDN:AAA5320.1; PID:g149854
A>Note: the authors translated the codon GCG for residue 331 as Pro

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog
C:Keywords: GTP binding; P-loop; protein biosynthesis

F:13-141/Domain: translation elongation factor Tu homology <ETU>
F:13-26/Region: nucleotide-binding motif A (P-loop)

F:136-141/Region: GTP-binding NKXD motif
F:176-178/Region: GTP-binding SAK/L motif
F:25,26,64,138,139,141,176/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #sta

Query Match
Best Local Similarity 82.9%; Score 34; DB 2; Length 396;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GRVLVQPG 8
Db 292 GQVLVVRG 299

RESULT 3

Probable protein-glutamate methyltransferase (cheb) - syphilis spirochete

C:Species: *Treponema pallidum* subsp. *pallidum* (syphilis spirochete)

C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 03-Dec-1999
C:Accession: E71300

R:Praser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khairak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998

A>Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770
A:Accession: E71300

A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-403 <COL>
A:Cross-references: GB:AE001238; GB:AE000520; NID:g3322928; PIDN:AAC65606.1; PID:g332293

A:Experimental source: strain Nichols
C:Genetics:

A:Gene: TP0631
C:Superfamily: protein-glutamate methyltransferase; response regulator homology
C:Keywords: phosphoprotein

F:9-122/Domain: response regulator homology <RRH>
F:39/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match
Best Local Similarity 82.9%; Score 34; DB 2; Length 403;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GRVLVQPG 8
Db 279 GRVLVQPG 286

RESULT 4

structural protein 2 precursor - hepatitis E virus (strain Mexico)

C:Species: hepatitis E virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: B44212

R:Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.
Virology 191, 550-558, 1992

A>Title: Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HE
A:Reference number: A44212; MUID:93079857
A:Accession: B44212

A:Molecule type: genomic RNA
A:Residues: 1-659 <HUA>

A:Cross-references: GB:M74506; NID:g330017; PIDN:AAA5732.1; PID:g330020
C:Superfamily: hepatitis E virus structural protein 2
C:Keywords: structural protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-659/Product: structural protein 2 #status predicted <SP2>

Query Match
Best Local Similarity 82.9%; Score 34; DB 1; Length 659;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 RVLVQPG 8
Db 233 RVLVQPG 239

RESULT 5

structural protein 2 precursor - hepatitis E virus (strain Burma)

C:Species: hepatitis E virus
C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 23-Jul-1999
C:Accession: C40778

R:Iam, A.W.; Smith, M.M.; Guerra, M.E.; Huang, C.C.; Bradley, D.W.; Fry, K.E.; Reyes,
Virology 185, 120-131, 1991

A>Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length
A:Reference number: A40778; MUID:92024067
A:Accession: C40778

A:Molecule type: genomic RNA
A:Residues: 1-660 <GRAM>

A:Cross-references: GB:M73218; NID:g330023; PIDN:AAA5736.1; PID:g330026
A>Note: the authors translated the codon GCG for residue 2 as Ala

C:Superfamily: hepatitis E virus structural protein 2
C:Keywords: structural protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-660/Product: structural protein 2 #status predicted <SP2>

Query Match
Best Local Similarity 82.9%; Score 34; DB 1; Length 660;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 RVLVQPG 8
Db 233 RVLVQPG 239

RESULT 6

hypothetical protein RV0236c - *Mycobacterium tuberculosis* (strain H37Rv)

C:Species: *Mycobacterium tuberculosis*
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 28-Jul-2000
C:Accession: B70963

R:Coile, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Rajandream, M.A.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: B70963

A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1400 <COL>

A:Cross-references: GB:Z92669; GB:AL123456; NID:g3242271; PIDN:CAB07017.1; PID:g32422
A:Experimental source: strain H37Rv
C:Genetics:

A:Gene: RV0236c
C:Superfamily: *Mycobacterium leprae* probable integral membrane protein

Query Match
Best Local Similarity 82.9%; Score 34; DB 2; Length 1400;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVOPG 8
11111111
DB 454 GRVLVOPG 461

RESULT 7

T44872
probable integral membrane protein [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 28-Jul-2000
C:Accession: T44872
R:Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1998
A:Reference number: 222863
A:Accession: T44872
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1440 <PAR>
A:Cross-references: EMBL:AL022486; PIDN:CA18562.1
A:Experimental source: cosmid B1883
C:Genetics:
A:Note: MLCB1883.13c
C:Superfamily: Mycobacterium leprae probable integral membrane protein

Query Match 82.9%; Score 34; DB 2; Length 1440;
Best Local Similarity 87.5%; Pred. No. 16+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVOPG 8
11111111
DB 489 GRVLVOPG 496

RESULT 8

B82694
copper homeostasis protein XFI341 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82694
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: For a complete list of authors see reference number A59328 below
A:Accession: B82694
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <SIM>
A:Cross-references: GB:AE003966; GB:AE003849; NID:g9106327; PIDN:AAF84150.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Slipson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Birones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Porty, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Latyr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sanceli, R.V.; Sawasak
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XFI341

Query Match 80.5%; Score 33; DB 2; Length 267;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVOPG 8
11111111
DB 194 GRVLVOPG 201

RESULT 9

T12497
hypothetical protein DKFP434K091.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999
C:Accession: T12497
R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z17525
A:Accession: T12497
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-321 <POU>
A:Cross-references: EMBL:AL080175
A:Experimental source: adult testis; clone DKFP434K091
C:Genetics:
A:Note: DKFP434K091.1

Query Match 80.5%; Score 33; DB 2; Length 321;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVLVOPG 8
11111111
DB 171 RVLVOPG 177

RESULT 10

G72643
probable transketolase APE0583 - Aeropyrum pernix (strain KI)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: G72643
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339
A:Accession: G72643
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-322 <KAW>
A:Cross-references: DBJ:AP000060; NID:g5104188; PIDN:BA79551.1; PID:d1043337; PID:g
A:Experimental source: strain KI
C:Genetics:
A:Gene: APE0583

Query Match 80.5%; Score 33; DB 2; Length 322;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVOPG 8
11111111
DB 189 GRVLVOPG 196

RESULT 11

E82850
fimbrial adhesin precursor XFO078 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: E82850
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: AB2515; MUID:20365717
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: E82850
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-354 <SIM>
 A:Cross-references: GB:AE003862; GB:AE003849; NID:g9104849; PIDN:AAF82891.1; GSPDB:GN001
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Relnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H
 as-Neto, E.; Docena, C.; El-Dorry, H.; Facincanli, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; From
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitaajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmiter, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Truhko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovskij-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0078

Query Match 80.5%; Score 33; DB 2; Length 354;
 Best Local Similarity 71.4%; Pred. No. 39;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GRVLVOP 7
 111111
 Db 55 GRVLVOP 61

RESULT 12
 S62726
 translation elongation factor tu (EF-tu) U0522 [similarity] - Ureaplasma urealyticum
 C:Species: Ureaplasma urealyticum
 C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 02-Sep-2000
 C:Accession: S62726; F82879
 R:Kamita, V.
 submitted to the EMBL Data Library, June 1994
 A:Reference number: S62726
 A:Accession: S62726
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-394 <KAM>
 A:Cross-references: EMBL:Z34275; NID:g498790; PIDN:CAA84029.1; PID:g498791
 R:Glass, J.I.; Lettkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.
 submitted to GenBank, February 2000
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
 A:Reference number: A82870
 A:Accession: F82879
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-394 <GLA>
 A:Cross-references: GB:AE002151; GB:AF222894; NID:g6699524; PIDN:AAF30935.1; GSPDB:GN001
 A:Experimental source: serovar 3; blover 1
 C:Genetics:
 A:Gene: tuf; U0522
 A:Genetic code: SGC3
 C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo
 C:Keywords: GTP binding; P-loop
 F:13-139/Domain: translation elongation factor Tu homology <ETU>
 F:19-26/Region: nucleotide-binding motif A (P-loop)
 F:136-139/Region: GTP-binding NKXD motif

Query Match 80.5%; Score 33; DB 2; Length 394;
 Best Local Similarity 75.0%; Pred. No. 43;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVLVOPG 8
 111111
 Db 290 GQVLVOPG 297

RESULT 13
 B70300
 translation elongation factor EF-Tu - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 26-Aug-1999
 C:Accession: B70300
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
 V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus..
 A:Reference number: A70300; MUID:98196666
 A:Accession: B70300
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-405 <AOF>
 A:Cross-references: GB:AE000669; NID:g2982762; PIDN:AAC06403.1; PID:g2982777; GB:AE00
 A:Experimental source: strain VF5
 C:Genetics:
 A:Gene: tufA1
 C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom
 C:Keywords: GTP binding; P-loop
 F:13-143/Domain: translation elongation factor Tu homology <ETU>
 F:19-26/Region: nucleotide-binding motif A (P-loop)
 F:140-143/Region: GTP-binding NKXD motif
 F:178-180/Region: GTP-binding SAK/L motif
 F:25,26,66,140,141,143,178/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #

Query Match 80.5%; Score 33; DB 2; Length 405;
 Best Local Similarity 75.0%; Pred. No. 45;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRVLVOPG 8
 111111
 Db 300 GQVLVOPG 307

RESULT 14
 G70465
 translation elongation factor EF-Tu - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 26-Aug-1999
 C:Accession: G70465
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
 V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus..
 A:Reference number: A70300; MUID:98196666
 A:Accession: G70465
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-405 <AOF>
 A:Cross-references: GB:AE000763; NID:g2984178; PIDN:AAC07714.1; PID:g2984182; GB:AE00
 A:Experimental source: strain VF5
 C:Genetics:
 A:Gene: tufA2
 C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom
 C:Keywords: GTP binding; P-loop
 F:13-143/Domain: translation elongation factor Tu homology <ETU>
 F:19-26/Region: nucleotide-binding motif A (P-loop)

Query Match 80.5%; Score 33; DB 2; Length 405;
 Best Local Similarity 75.0%; Pred. No. 45;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRVLVOPG 8
 111111

Db 300 GOVLA0PG 307

RESULT 15

G69256

conserved hypothetical protein AF0055 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C:Accession: G69256

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343

A:Accession: G69256

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1132 <KLE>

A:Cross-references: GB:AE001102; GB:AE000782; MID:g2689425; PIDN:AAB9116.1; PID:g265058

C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1407

Query Match

78.0%; Score 32; DB 1; Length 132;

Best Local Similarity 75.0%; Pred. NO. 23;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GRVLVQPG 8

Db 50 GRILVFPG 57

Search completed: June 13, 2001, 14:10:38
Job time: 139 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:21:39 ; Search time 74.44 Seconds
(without alignments)
3.681 Million cell updates/sec

Title: PCT-US01-05825a-3

Perfect score: 41

Sequence: 1 GRVLVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	82.9	396	1	EFTU_MICLU
2	34	82.9	403	1	CHEB_TREPA
3	34	82.9	485	1	VST2_HEVRH
4	34	82.9	659	1	VST2_HEVME
5	34	82.9	660	1	VST2_HEVBU
6	34	82.9	660	1	VST2_HEVMA
7	34	82.9	660	1	VST2_HEVPA
8	34	82.9	713	1	GAC1_HUMAN
9	34	82.9	836	1	GLK1_MOUSE
10	34	82.9	1122	1	TERT_MOUSE
11	33	80.5	394	1	EFTU_TREPA
12	33	80.5	405	1	EFTU_AQUAE
13	33	80.5	405	1	EFTU_AQUAE
14	33	80.5	405	1	EFTU_AQUAE
15	32	78.0	132	1	Y055_ARCFU
16	32	78.0	440	1	LCAT_RABIT
17	32	78.0	461	1	HGD_ARATH
18	32	78.0	569	1	GCL1_DROME
19	31	75.6	235	1	EFTU_GRALE
20	31	75.6	367	1	INX2_DROME
21	31	75.6	396	1	EFTU_CORGL
22	31	75.6	397	1	EFTU_BREIN
23	31	75.6	460	1	Y042_CAEEL
24	31	75.6	2225	1	Y042_CAEEL
25	31	75.6	2225	1	Y042_CAEEL
26	31	75.6	2225	1	Y042_CAEEL
27	30	73.2	116	1	NS2C_HUMAN
28	30	73.2	353	1	NS2C_HUMAN
29	30	73.2	355	1	NS2C_HUMAN
30	30	73.2	381	1	NS2C_HUMAN
31	30	73.2	387	1	NS2C_HUMAN
32	30	73.2	397	1	NS2C_HUMAN
33	30	73.2	397	1	NS2C_HUMAN

ALIGNMENTS

34	30	73.2	397	1	EFT1_STRRA	P29542 streptomyc
35	30	73.2	397	1	EFTU_STRAU	O33594 streptomyc
36	30	73.2	548	1	G01M_ECOLI	P23714 escherichia
37	30	73.2	736	1	GEPH_RAT	Q03555 rattus norv
38	30	73.2	916	1	TOP1_ARATH	P30181 arabidopsis
39	30	73.2	1318	1	VP14_EBV	P03179 Epstein-Barr
40	30	73.2	1331	1	XDH_BOVIN	P80457 bos taurus
41	30	73.2	1379	1	MDH_MOUSE	P16056 mus musculu
42	30	73.2	1390	1	MDH_MOUSE	P08581 mus musculu
43	30	73.2	1453	1	NKCR_MOUSE	P30415 mus musculu
44	30	73.2	4344	1	DYHC_EMENT	P45444 emericella
45	30	73.2	4349	1	DYHC_FUSSO	P78716 fusarium so

RESULT 1
EFTU_MICLU STANDARD: PRT: 396 AA.
AC P09953;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ELONGATION FACTOR TU (EF-TU).
GN TUF OR TUF.
OS Micrococcus luteus (Micrococcus lysodeikticus).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Micrococcus.
OX NCBI_TaxID=1270;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8007427; PubMed=3654584;
RA Ohana T., Yamao F., Muto A., Osawa S.;
RT "Organization and codon usage of the streptomycin operon in
Micrococcus luteus, a bacterium with a high genomic G + C content.";
RL J. Bacteriol. 169:4770-4777(1987)
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M17788; AAA25320.1; -.
CC PIR: D26956; D26956.
CC DR HSP: P02990; 1EFTU.
CC DR InterPro: IPR000795; -.
CC DR Pfam: PF00009; GTP_EFTU; 1.
CC DR PRINTS: PR00315; ELONGATNFCP.
CC DR PROSITE: PS00301; EFACOR_GTP; 1.
CC KW Elongation factor; protein biosynthesis; GTP-binding.
CC FT NP_BIND 19 26 GTP (BY SIMILARITY).
CC FT NP_BIND 83 87 GTP (BY SIMILARITY).
CC FT NP_BIND 138 141 GTP (BY SIMILARITY).
CC SQ SEQUENCE 396 AA; 43820 MW; 5285697D5E3F2359 CRC64;
Query Match 82.9%; Score 34; DB 1; Length 396;
Best Local Similarity 75.0%; Pred. No. 9.1;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
1 GRVLVQPG 8
|:|||||

Db 292 GQVLEPG 299

```

RESULT 2
CHEB_TREPA
ID CHEB_TREPA STANDARD: PRT: 403 AA.
AC 083639:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE PROTEIN-GLUTAMATE METHYLESTERASE (EC 3.1.1.61).
GN CHEB OR TP0631.
OS Treponema pallidum.
OC Bacteria: Spirochaetales: Spirochaetaceae: Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RA MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artlach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -1- FUNCTION: INVOLVED IN THE MODULATION OF THE CHEMOTAXIS SYSTEM;
CC CATALYZES THE DEMETHYLATION OF SPECIFIC METHYLGLUTAMATE RESIDUES
CC INTRODUCED INTO THE CHEMORECEPTORS (METHYL-ACCEPTING CHEMOTAXIS
CC PROTEINS) BY CHER (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: PROTEIN L-GLUTAMATE O-METHYL ESTER + H(2)O =
CC PROTEIN L-GLUTAMATE + METHANOL.
CC -1- SUBCELLULAR LOCATION: CYTOSOLASMIC.
CC -1- DOMAIN: THE N-TERMINAL REGULATORY DOMAIN INHIBITS THE ACTIVITY OF
CC THE C-TERMINAL EFFECTOR DOMAIN.
CC -1- PHM: PHOSPHORYLATED BY CHEA. PHOSPHORYLATION SUPPRESSES THE
CC INHIBITORY ACTIVITY OF THE N-TERMINAL DOMAIN (BY SIMILARITY).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION: BELONGS TO THE RESPONSE
CC REGULATORY FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION: BELONGS TO THE CHEB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001238; AAC65606.1; -.
DR HSSP: P04042; ICND.
DR TIGR: TP0631; -.
DR InterPro: IPR000673; -.
DR InterPro: IPR001789; -.
DR Pfam: PF01339; Cheb_methylase; 1.
DR Pfam: PF00072; response_reg; 1.
KM Hydrolyase; Chemotaxis; Sensory transduction; Phosphorylation.
FT DOMAIN 1 138 RESPONSE REGULATORY DOMAIN.
FT DOMAIN 139 202 LINKER.
FT DOMAIN 203 403 PROTEIN-GLUTAMATE METHYLESTERASE.
FT MOD_RES 59 59 PHOSPHORYLATION (BY SIMILARITY).
FT ACT_SITE 219 219 BY SIMILARITY.
FT ACT_SITE 246 246 BY SIMILARITY.
FT ACT_SITE 342 342 BY SIMILARITY.
SQ SEQUENCE 403 AA; 44098 MW; E629B147BF3AD03A CRC64;

```

Query Match 82.9%; Score 34; DB 1; Length 403;
 Best Local Similarity 75.0%; Pred. No. 9.2;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVPG 8
 Db 279 GRVLVPG 286

```

RESULT 3
VST2_HEVME
ID VST2_HEVME STANDARD: PRT: 485 AA.
AC 000270:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE STRUCTURAL PROTEIN 2 (FRAGMENT).
OS Hepatitis E virus (isolate Rhesus) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage.
OX NCBI_TaxID=31766;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K.,
RA Wain K.M.;
RT "Hepatitis E virus: cDNA cloning and expression."
RL Microbiol. Immunol. 36:67-79(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D90274; BAA20910.1; -.
FT NON_TER 1 1
FT NON_TER 485 485
SQ SEQUENCE 485 AA; 52317 MW; 5A4FD9273AC74F94 CRC64;

```

Query Match 82.9%; Score 34; DB 1; Length 485;
 Best Local Similarity 85.7%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 2 RVLVPG 8
Db 101 RVLVPG 107

RESULT 4
VST2_HEVME
ID VST2_HEVME STANDARD: PRT: 659 AA.
AC 003500:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE STRUCTURAL PROTEIN 2 PRECURSOR.
OS Hepatitis E virus (strain Mexico) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage.
OX NCBI_TaxID=31766;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93079857; PubMed=1448913;
RA Huang C.C., Nguyen D., Fernandez J., Yun K.Y., Fry K.E.,
RA Bradley D.W., Tam A.W., Reyes G.R.;
RT "Molecular cloning and sequencing of the Mexico isolate of hepatitis
RT E virus (HEV)."
RL Virology 191:550-558(1992).
CC -1- FUNCTION: CONTRAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M74506; AAA45732.1; -
DR PIR: B44212; B44212.
KW Signal.
FT CHAIN 1 22 BY SIMILARITY.
SO SEQUENCE 659 AA; 70640 MW; CP75E75EPD8FB2C CRC64;

Query Match
Best Local Similarity 82.9%; Score 34; DB 1; Length 659;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVLVOPG 8
DB 233 RILVOPG 239

RESULT 5
VST2_HEVBU STANDARD; PRT: 660 AA.
AC P39326;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE STRUCTURAL PROTEIN 2 PRECURSOR (ORF2).
OS Hepatitis E virus (strain Burma) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage.
OX NCBI_Taxid=31767;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92024067; PubMed-1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M73218; AAA45736.1; -
DR PIR: C40778; VHWHR2.
KW Signal.
FT CHAIN 1 19 POTENTIAL.
FT SIGNAL 20 660 STRUCTURAL PROTEIN 2.
SO SEQUENCE 660 AA; 70978 MW; 5832A013CCCA61C CRC64;

Query Match
Best Local Similarity 82.9%; Score 34; DB 1; Length 660;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVLVOPG 8
DB 233 RILVOPG 239

RESULT 6

```

```

VST2_HEVWY STANDARD; PRT: 660 AA.
AC 004611;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE STRUCTURAL PROTEIN 2 PRECURSOR (ORF2)
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage.
OX NCBI_Taxid=31769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93227573; PubMed-8470371;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikhsa T., Winn K.;
RT "Sequence and gene structure of the hepatitis E virus isolated from
RT Myanmar.";
RL Virus Genes 7:95-109(1993).
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D10330; BAA01174.1; -
DR PIR: D10330; BAA01174.1; -
KW Signal.
FT CHAIN 1 22 BY SIMILARITY.
FT SIGNAL 23 660 STRUCTURAL PROTEIN 2.
SO SEQUENCE 660 AA; 70998 MW; 3A82A4EA255C6253 CRC64;

Query Match
Best Local Similarity 82.9%; Score 34; DB 1; Length 660;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVLVOPG 8
DB 233 RILVOPG 239

RESULT 7
VST2_HEVPA STANDARD; PRT: 660 AA.
AC P33426;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE STRUCTURAL PROTEIN 2 PRECURSOR (ORF2).
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage.
OX NCBI_Taxid=33774;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92115700; PubMed-1713327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik T.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: M80581; AAA45727.1; -
 KW SIGNAL.
 FT SIGNAL. 1 22 BY SIMILARITY.
 FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
 SO SEQUENCE 660 AA; 70980 MW; 8085BC53CFB46FD3 CRC64;

Query Match
 Best Local Similarity 82.9%; Score 34; DB 1; Length 660;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVLVOPG 8
 1:|||||
 Db 233 RILVOPG 239

RESULT 8

GAC1_HUMAN STANDARD; PRT; 713 AA.

AC 075325;
 DT 01-OCT-2000 (Rel. 40; Created)
 DT 01-OCT-2000 (Rel. 40; Last sequence update)
 DT 01-OCT-2000 (Rel. 40; Last annotation update)
 DE GLIOMA AMPLIFIED ON CHROMOSOME 1 PROTEIN PRECURSOR.

GN GAC1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxId=9606;
 RN [1]

RP TISSUE: Glioma tumor;
 RC MEDLINE=96324709; PubMed=9662332;

RA Malfroy B., Almeida A., Zhu X.X., Vogt N., Tyagi R., Mulieris M.,
 RA Dutrillaux A.-M., Dutrillaux B., Ross D., Hansh S.,
 RT "GAC1, a new member of the leucine-rich repeat superfamily on
 RT chromosome band 1q32.1, is amplified and overexpressed in malignant
 RT gliomas."

RL Oncogene 16:2997-3002(1998).

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).

CC -1- TISSUE SPECIFICITY: OVERAMPLIFIED IN MALIGNANT GLIOMAS.

CC -1- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).

CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AF030435; AAC39792.1; -

DR MIM: 605492; -

DR InterPro: IPR000372; -

DR InterPro: IPR000483; -

DR InterPro: IPR001611; -

DR InterPro: IPR003006; -

DR Pfam: PF00047; 19; 1.

DR Pfam: PF00560; LRR; 9.

DR Pfam: PF01462; LRRNT; 1.

DR Pfam: PF01463; LRRCT; 1.

DR PRINTS: PR00019; LEURCHRP.

DR Immunoglobulin domain; Transmembrane; Glycoprotein; Repeat;

KW Leucine-rich repeat; Signal.

FT SIGNAL. 1 18

FT CHAIN 19 713

FT DOMAIN 19 630

FT TRANSMEM 631 651

FT DOMAIN 652 713

CC POTENTIAL.
 CC GLIOMA AMPLIFIED ON CHROMOSOME 1 PROTEIN.
 CC EXTRACELLULAR (POTENTIAL).
 CC POTENTIAL.
 CC CYTOPLASMIC (POTENTIAL).

FT REPEAT 92 115 LRR 1.
 FT REPEAT 116 139 LRR 2.
 FT REPEAT 140 163 LRR 3.
 FT REPEAT 165 187 LRR 4.
 FT REPEAT 188 211 LRR 5.
 FT REPEAT 213 235 LRR 6.
 FT REPEAT 236 259 LRR 7.
 FT REPEAT 261 283 LRR 8.
 FT REPEAT 309 333 LRR 9.
 FT REPEAT 334 357 LRR 10.
 FT REPEAT 359 385 LRR 11.
 FT DOMAIN 438 504
 FT DIISULFID 445 497
 FT CARBOHYD 94
 FT CARBOHYD 381
 FT CARBOHYD 381
 FT CARBOHYD 555
 FT CARBOHYD 583
 SO SEQUENCE 713 AA; 78798 MW; EC8BC0FD240C9396 CRC64;

Query Match
 Best Local Similarity 82.9%; Score 34; DB 1; Length 713;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVLVOPG 8
 1:|||||
 Db 515 GRALLVOPG 522

RESULT 9

GLK1_MOUSE STANDARD; PRT; 836 AA.

AC 060934;
 DT 15-JUL-1998 (Rel. 36; Created)
 DT 15-JUL-1998 (Rel. 36; Last sequence update)
 DT 01-OCT-2000 (Rel. 40; Last annotation update)
 DE GLUTAMATE RECEPTOR, IONOTROPIC KAINATE 1 PRECURSOR (GLUTAMATE RECEPTOR
 DE 5) (GLUR-5).
 GN GRIK1 OR GLUR5.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=Brain;

RA MEDLINE=94083547; PubMed=8260617;

RT Gregor P., O'Hara B.F., Yang X., Uhl G.R.,

RT "Expression and novel subunit isoforms of glutamate receptor genes

RT Glur5 and Glur6."

RT NeuroReport 4:1343-1346(1993).

RN [2]

RP SEQUENCE OF 584-695 FROM N.A., AND RNA EDITING.

RC STRAIN=BALB/C;

RA MEDLINE=96312506; PubMed=8700852;

RT Herb A., Higuchi M., Sprengel R., Seeburg P.H.;

RT "Q/R site editing in kainate receptor Glur5 and Glur6 pre-mRNAs

RT requires distant intronic sequences."

RT Proc. Natl. Acad. Sci. U.S.A. 93:1875-1880(1996).

RN [3]

RP SEQUENCE OF 606-622 FROM N.A., AND RNA EDITING.

RC MEDLINE=92005683; PubMed=1717158;

RA Sommer B., Koehler M., Sprengel R., Seeburg P.H.;

RT "RNA editing in brain controls a determinant of ion flow in

RT glutamate-gated channels."

RT Cell 67:11-19(1991).

CC -1- FUNCTION: L-GLUTAMATE ACTS AS AN EXCITATORY NEUROTRANSMITTER AT

CC ACTIONS OF GLU ARE MEDIATED BY A VARIETY OF RECEPTORS THAT ARE

CC NAMED ACCORDING TO THEIR SELECTIVE AGONISTS. MAY BE INVOLVED IN

CC THE TRANSMISSION OF LIGHT INFORMATION FROM THE RETINA TO THE

CC HYPOTHALAMUS (BY SIMILARITY).

CC -1- SUBUNIT: THE UNEDITED VERSION OF (Q) ASSEMBLES INTO A FUNCTIONAL

CC KAINATE-GATED HOMOMERIC CHANNEL, WHEREAS THE EDITED VERSION (R) IS
 CC UNABLE TO PRODUCE CHANNEL ACTIVITY WHEN EXPRESSED ALONE. BOTH
 CC EDITED AND UNEDITED VERSIONS CAN FORM FUNCTIONAL CHANNELS WITH
 CC GRIK4 AND GRIK5 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN THE CEREBELLUM. ALSO PRESENT
 CC IN THE SUPRACHIASMATIC NUCLEI OF THE HYPOTHALAMUS.
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X66118; NOT_ANNOTATED_CDS.
 DR EMBL: U31444; AA083222.1; -.
 DR MGI: 95814; Grik1.
 DR InterPro: IPR001320; -.
 DR InterPro: IPR001828; -.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR Pfam: PF00060; lig_chan; 1.
 DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Phosphorylation; Alternative splicing; RNA editing.
 FT SIGNAL 1 30
 FT CHAIN 31 836
 FT DOMAIN 31 561
 FT TRANSMEM 562 582
 FT TRANSMEM 639 659
 FT TRANSMEM 722 742
 FT CARBOHYD 68 68
 FT CARBOHYD 74 74
 FT CARBOHYD 276 276
 FT CARBOHYD 379 379
 FT CARBOHYD 413 413
 FT CARBOHYD 424 424
 FT CARBOHYD 431 431
 FT CARBOHYD 546 546
 FT VARIANT 621 621
 FT SEQUENCE 836 AA; 95201 MW; ABB6EC4B44FFDBFF CRC64;
 SO SEQUENCE

Query Match 82.9%; Score 34; DB 1; Length 836;
 Best Local Similarity 75.0%; Pred. No. 19;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVQPG 8
 DB 4 GTVLVQPG 11

RESULT 10
 TERT_MOUSE
 ID TERT_MOUSE STANDARD: PRT; 1122 AA.
 AC 070372; 035432;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE TELOMERASE REVERSE TRANSCRIPTASE (EC 2.7.7.-) (TELOMERASE CATALYTIC
 DE SUBUNIT).
 GN TERT.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98241176; PubMed=9582020;
 RA Greenberg R.A., Allsopp R.C., Chin L., Morin G.B., DePinho R.A.;

RT "Expression of mouse telomerase reverse transcriptase during
 RT development, differentiation and proliferation."
 RT Oncogene 16:1723-1730(1998).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98393668; PubMed=9724727;
 RA Martin-Rivera L., Herrera E., Albar J.P., Blasco M.A.
 RT "Expression of mouse telomerase catalytic subunit in embryos and
 RT adult tissues."
 RT Proc. Natl. Acad. Sci. U.S.A. 95:10471-10476(1998).
 RN [3]
 RP SEQUENCE OF 550-616 FROM N.A.
 RA Drissi R., Cleveland J.L.,
 RT "Partial sequence of Mus musculus telomerase catalytic subunit
 RT homolog.";
 RT Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR
 CC THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT
 CC ELONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS
 CC SIMPLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE
 CC SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.
 CC TELOMERASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF051911; AAC09323.1; -.
 DR EMBL: AF073311; AAC34821.1; -.
 DR EMBL: AF029235; AAB84200.1; -.
 DR MGI: 1202709; Tert.
 KW Transferrase; RNA-directed DNA polymerase; Telomere; Nuclear protein;
 KW DNA-binding.
 FT CONFLICT 553 553
 FT SEQUENCE 1122 AA; 127977 MW; F8526695DD6558C CRC64;
 SO SEQUENCE

Query Match 82.9%; Score 34; DB 1; Length 1122;
 Best Local Similarity 87.5%; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVQPG 8
 DB 35 GRRVQPG 42

RESULT 11
 EFTU_UREPA
 ID EFTU_UREPA STANDARD: PRT; 394 AA.
 AC P50068;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ELONGATION FACTOR TU (EF-TU).
 GN TUF OR UUS522.
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscutes;
 OC Mycoplasmataceae; Ureaplasma.
 OX NCBI_TaxID=134821;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=ATCC 33697 / SEROVAR 14;
 RA Bruex A.;
 RL Thesis (1994), Heinrich-Heine University / Duesseldorf, Germany.
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=SEROVAR 3;

```

RA MEDLINE-20500219; PubMed-11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Casseil G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum."
RL Nature 407:757-762(2000).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 234275; CAA84029.1;
DR EMBL; AE002151; AAF30935.1;
DR HSSP; P02990; IFTU
DR InterPro: IPR000795;
DR Pfam: PF00009; GTP_EFTU; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFACOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
FT SEQUENCE 394 AA; 42902 MW; 80A887B6C59883E0 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 394;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVLVOPG 8
DB 290 GQVLVPG 297

RESULT 12
EFTU_AOUAE STANDARD; PRT; 405 AA.
AC 066428; 067755;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ELONGATION FACTOR TU (EF-TU).
GN (TUF1 OR AQ.005) AND (TUPA2 OR AQ.1928).
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VFS;
RX MEDLINE-98196666; PubMed-9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Shead M.A., Kellner M., Anjay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Sanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998)
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.

```

```

CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000669; AAC06403.1;
DR EMBL; AE000763; AAC07714.1;
DR InterPro: IPR000795;
DR Pfam: PF00009; GTP_EFTU; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFACOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding;
KW Multigene family.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 85 89 GTP (BY SIMILARITY).
FT NP_BIND 140 143 GTP (BY SIMILARITY).
FT VARIANT 140 140 N -> S (IN TUPA2).
FT VARIANT 294 294 K -> R (IN TUPA2).
FT SEQUENCE 405 AA; 44743 MW; FBE083116BE15F1D CRC64;

Query Match 80.5%; Score 33; DB 1; Length 405;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVLVOPG 8
DB 300 GQVLVPG 307

RESULT 13
EFTU_AOUAY STANDARD; PRT; 405 AA.
AC 050293;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ELONGATION FACTOR TU (EF-TU).
GN TUF.
OS Aquifex pyrophilus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=2714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 6858;
RX MEDLINE-98248216; PubMed-9588802;
RA Ludwig W., Strunk O., Klugbauer S., Klugbauer N., Weizenegger M.,
RA Neumaier J., Bachleitner M., Schleifer K.H.;
RT "Bacterial phylogeny based on comparative sequence analysis."
RL Electrophoresis 19:554-568(1998)
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y15787; CAA75781.1;
DR InterPro: IPR000795;

```

DR Pfam: PF00009: GTP_EFTU: 1.
DR PRINTS: PR00315; ELONGATNCT.
DR PROSITE: PS00301; EFATOR_GTP: 1.
KW Elongation factor; Protein biosynthesis; GTP-binding.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 85 89 GTP (BY SIMILARITY).
FT NP_BIND 140 143 GTP (BY SIMILARITY).
SO SEQUENCE 405 AA: 44672 MW: 012A9CD9C60AA16C CRC64;

Query Match 80.5%; Score 33; DB 1; Length 405;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVOPG 8
DB 300 GOVLAQPG 307

RESULT 14
Y055_ARCFU STANDARD: PRT: 132 AA.
ID Y055_ARCFU
AC Y03181;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL PROTEIN AF0055.
GN AF0055.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klank H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyriides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kikunas E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997)
CC -1- SIMILARITY: BELONGS TO THE UPF0107 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE001102; AAB91166.1; -
DR TIGR: AF0055; -
DR InterPro: IPR002840; -
DR Pfam: PF01989; DUF126; 1.
KW Hypothetical protein.
SO SEQUENCE 132 AA: 14189 MW: 9C97ABCF9512F20D CRC64;

Query Match 78.0%; Score 32; DB 1; Length 132;
Best Local Similarity 75.0%; Pred. No. 8.1;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVOPG 8

DB 50 GRVLVOPG 57
ID LCAT_RABIT
AC P53761.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PHOSPHATIDYLCHOLINE-STEROL ACYLTRANSFERASE PRECURSOR (EC 2.3.1.43)
DE (LECITHIN-CHOLESTEROL ACYLTRANSFERASE) (PHOSPHOLIPID-CHOLESTEROL
DE ACYLTRANSFERASE).
GN LCAT.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND;
RX MEDLINE=96425081; PubMed=8827532;
RA Murata Y., Maeda E., Yoshino G., Kasuga M.;
RT "Cloning of rabbit LCAT cDNA: increase in LCAT mRNA abundance in the
RT liver of cholesterol-fed rabbits."
RL J. Lipid Res. 37:1616-1622(1996).
CC -1- FUNCTION: CENTRAL ENZYME IN THE EXTRACELLULAR METABOLISM OF PLASMA
CC LIPOPROTEINS. AMONG OTHER SUBSTRATES IT ESTERIFIES THE FREE
CC CHOLESTEROL. TRANSPORTED IN PLASMA LIPOPROTEINS.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + STEROL = STEROL ESTER +
CC -1- ACYLGLYCEROPHOSPHOCHOLINE (PALMITOYL, OLEOYL, AND LINOLEOYL CAN
CC BE TRANSFERRED; A NUMBER OF STEROLS, INCLUDING CHOLESTEROL, CAN
CC ACT AS ACCEPTOR).
CC -1- ENZYME REGULATION: APOLIPOPROTEIN A-I IS A POTENT ACTIVATOR FOR
CC THIS ENZYME.
CC -1- SIMILARITY: PARTIAL WITH LIPASES (PANCREATIC, GASTRIC, HEPATIC,
CC LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D13668; BAA02839.1; -
DR InterPro: IPR000734; -
DR PROSITE: PS00120; LIPASE_SER; 1.
KW transferase; Acyltransferase; Lipid metabolism; Glycoprotein; Signal.
FT FT SIGNAL 1 24
FT CHAIN 25 440
FT FT
FT ACT_SITE 205 205
FT DISULFID 74 98
FT FT DISULFID 337 380
FT CARBOHYD 44 44
FT CARBOHYD 108 108
FT CARBOHYD 296 296
FT CARBOHYD 408 408
SO SEQUENCE 440 AA: 49559 MW: 1958C5B43BD534AD CRC64;

Query Match 78.0%; Score 32; DB 1; Length 440;
Best Local Similarity 62.5%; Pred. No. 26;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVOPG 8
DB 112 GRVLVOPG 119

Wed Jun 13 15:00:49 2001

pct-us01-05825a-3.rsp

Search completed: June 13, 2001, 14:21:40
Job time: 800 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:20:22 ; Search time 225.85 Seconds
(Without alignments)
4.152 Million cell updates/sec

Title: PCT-US01-05825A-3
Perfect score: 41
Sequence: 1 GRVLPQG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.15:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	90.2	399	5	Q9NGF2
2	37	90.2	399	5	Q9NGA6
3	37	90.2	347	4	Q9S072
4	37	90.2	547	4	Q9NVQ9
5	37	90.2	2189	5	Q9VXD5
6	35	85.4	1128	11	Q9QXZ4
7	34	82.9	104	11	Q9JLM1
8	34	82.9	227	14	Q9S048
9	34	82.9	227	14	Q9S049
10	34	82.9	227	14	Q9S049
11	34	82.9	253	2	Q9PEW8
12	34	82.9	283	14	Q9YBP1
13	34	82.9	436	14	Q9W148
14	34	82.9	466	14	Q9WNN1
15	34	82.9	486	14	Q9WNN2
16	34	82.9	514	11	Q9R266
17	34	82.9	525	14	Q9Y947
18	34	82.9	605	14	Q9YWL0
19	34	82.9	660	14	Q69419

20	34	82.9	660	14	Q69411	069411 hepatitis e
21	34	82.9	660	14	Q81871	Q81871 hepatitis e
22	34	82.9	660	14	Q89468	Q89468 hepatitis e
23	34	82.9	660	14	Q71147	Q71147 hepatitis e
24	34	82.9	660	14	Q68985	Q68985 hepatitis e
25	34	82.9	660	14	Q91855	Q91855 hepatitis e
26	34	82.9	660	14	Q91856	Q91856 hepatitis e
27	34	82.9	660	14	Q81878	Q81878 hepatitis e
28	34	82.9	660	14	Q36613	Q36613 swine hepat
29	34	82.9	660	14	Q9YLR2	Q9YLR2 hepatitis e
30	34	82.9	660	14	Q9YLQ9	Q9YLQ9 hepatitis e
31	34	82.9	660	14	Q9WQAO	Q9WQAO hepatitis e
32	34	82.9	660	14	Q9WLL4	Q9WLL4 hepatitis e
33	34	82.9	660	14	Q9W7W9	Q9W7W9 hepatitis e
34	34	82.9	672	14	Q91V28	Q91V28 hepatitis e
35	34	82.9	713	4	Q75325	Q75325 homo sapien
36	34	82.9	1400	2	P96419	P96419 mycobacteri
37	34	82.9	1440	2	Q69498	Q69498 mycobacteri
38	33	80.5	117	2	Q9Z5X9	Q9Z5X9 frankia sp.
39	33	80.5	267	2	Q9PDN8	Q9PDN8 xyella fas
40	33	80.5	321	4	Q9Y4N7	Q9Y4N7 homo sapien
41	33	80.5	322	1	Q9YEJ5	Q9YEJ5 aeropyrum p
42	33	80.5	354	2	Q9PH68	Q9PH68 xyella fas
43	33	80.5	392	5	Q9VMZ3	Q9VMZ3 drosophila
44	33	80.5	623	4	Q9P2V4	Q9P2V4 homo sapien
45	32	78.0	273	4	Q9Y321	Q9Y321 homo sapien

ALIGNMENTS

RESULT 1	Q9NGF2	PRELIMINARY;	PRT;	399 AA.
ID	Q9NGF2			
AC	Q9NGF2			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	RUDIMENTARY (FRAGMENT).			
GN	R.			
OS	Drosophila simulans (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7240;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-SIM3;			
RA	Begun D.J., Whitley P.;			
RT	"Reduced X-linked nucleotide polymorphism in Drosophila simulans.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 97:5960-5965(2000).			
DR	EMBL; AF252765; AAF68548.1; -			
FT	NON_TER			
FT	NON_TER			
SQ	SEQUENCE 399 AA; 44812 MW; 21ECDA04DA1A7CF0 CRC64;			
Query Match	90.2%; Score 37; DB 5; Length 399;			
Best Local Similarity	87.5%; Pred. No. 15;			
Matches	7; Conservative	1; Mismatches	0; Indels	0;
QY	1 GRVLPQG 8			
DB	256 GGVLPQG 263			
RESULT 2	Q9NGA6	PRELIMINARY;	PRT;	399 AA.
AC	Q9NGA6			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			

```

DE  RUDIMENTARY (FRAGMENT).
OS  Drosophila simulans (fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxID=7240;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=SIM8, SIM4, SIM5, SIM6, AND SIM7;
RA  Begun D.J., Whitely P.;
RT  "Reduced X-linked nucleotide polymorphism in Drosophila simulans.";
RL  Proc. Natl. Acad. Sci. U.S.A. 97:5960-5965(2000).
DR  EMBL: AF252770; AAF68553.1; -
DR  EMBL: AF252766; AAF68549.1; -
DR  EMBL: AF252767; AAF68550.1; -
DR  EMBL: AF252768; AAF68551.1; -
DR  EMBL: AF252769; AAF68552.1; -
FT  NON_TER 1
FT  SEQUENCE 399 AA; 44759 MW; AAF68552.1; CRC64;
SQ

```

```

Query Match
Best Local Similarity 90.2%; Score 37; DB 5; Length 399;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 GRVLYOPG 8
DB 256 GRVLYOPG 263

```

```

RESULT 3
ID 095072 PRELIMINARY; PRT; 547 AA.
AC 095072;
DT 01-MAY-1999 (TREMBLrel. 10. Created)
DT 01-MAY-1999 (TREMBLrel. 10. Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14. Last annotation update)
DE HHR21SPB (FRAGMENT).
GN HHR21SPB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA McKay M.J., van der Spek P., Kanaar R., Bootsma D., Hoeijmakers J.H.;
RT "HHR21SPB, a second human gene homologous to the rad21
RT Schizosaccharomyces pombe DNA double strand break repair gene.";
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF006264; AAD01193.1; -
FT NON_TER 547
FT SEQUENCE 547 AA; 62613 MW; 544AA31F2D03BD7B CRC64;
SQ

```

```

Query Match
Best Local Similarity 90.2%; Score 37; DB 4; Length 547;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 GRVLYOPG 8
DB 536 GRVLYOPG 543

```

```

DE  CDNA FLJ10573 FIS, CLONE NT2RP2003177.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA  Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA  Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA  Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA  Wakamatsu A., Nakamura T., Nagamori K., Masuno Y., Oshima A.;
RT  "NEDO human cDNA sequencing project.";
RL  Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR  EMBL: AK001435; BAA91690.1; -
FT  SEQUENCE 547 AA; 62585 MW; 05A1CF874F6C3BB CRC64;
SQ

```

```

Query Match
Best Local Similarity 90.2%; Score 37; DB 4; Length 547;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 GRVLYOPG 8
DB 536 GRVLYOPG 543

```

```

RESULT 5
ID 09VXD5 PRELIMINARY; PRT; 2189 AA.
AC 09VXD5;
DT 01-MAY-2000 (TREMBLrel. 13. Created)
DT 01-MAY-2000 (TREMBLrel. 13. Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15. Last annotation update)
DE R GENE PRODUCT.
GN R.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Broxstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Paulis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod K.P., McPherson D.,
RA Merkulov G., Mishaeva N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitsky R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveloff J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.W., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL: AEO03503; AAF48639.1; -
DR HSSP: P00479; 3CSU.
DR FLYBASE: FBgn0003189; r.
DR INTERPRO: IPR000901; -
DR INTERPRO: IPR000991; -
DR INTERPRO: IPR001317; -
DR INTERPRO: IPR002029; -
DR INTERPRO: IPR002082; -
DR INTERPRO: IPR002195; -
DR INTERPRO: IPR002385; -
DR INTERPRO: IPR002474; -
DR PFAM: PF00117; GATase; 1.
DR PFAM: PF00185; OTCase; 1.
DR PFAM: PF00289; CPSase_L_chain; 2.
DR PFAM: PF00744; Dihydroorotase; 1.
DR PFAM: PF00988; CPSase_sm_chain; 1.
DR PRINTS: PR00096; GATase.
DR PRINTS: PR00097; ANTSNTHASEII.
DR PRINTS: PR00098; CPSase.
DR PRINTS: PR00099; CPSGATase.
DR PRINTS: PR00100; AOTCase.
DR PRINTS: PR00101; ATCase.
DR PROSITE: PS00097; CARBAMOYLTRANSFERASE; UNKNOWN_1.
DR PROSITE: PS00442; GATase_TYPE_1; 1.
DR PROSITE: PS00482; DIHYDROOROTASE_1; 1.
DR PROSITE: PS00483; DIHYDROOROTASE_2; 1.
DR PROSITE: PS00866; CPSase_1; 1.
DR PROSITE: PS00867; CPSase_2; 2.
SQ SEQUENCE 2189 AA; 242759 MW; 42FDC09A6136A94F CRC64;

Query Match 90.2%; Score 37; DB 5; Length 2189;
Best Local Similarity 87.5%; Pred. No. 88;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLVQPG 8
DB 1777 GOVLVQPG 1784
1:|||||

RESULT 6
O90XZ4 PRELIMINARY; PRT; 1128 AA.
AC O90XZ4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE TELOMERASE CATALYTIC SUBUNIT.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
CC NCBI_TaxID=10036;
RN NCBI_TaxID=10036;
RP SEQUENCE FROM N.A.
RA Guo W., Okamoto M., Baluda M.A., Park N.-H.;
RT "Cloning of the cDNA and promoter of hamster telomerase catalytic
RT subunit (hamTERT).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF19012; AAF17334.1; -
SQ SEQUENCE 1128 AA; 128393 MW; 1DAF81249012174E CRC64;

Query Match 85.4%; Score 35; DB 11; Length 1128;
Best Local Similarity 87.5%; Pred. No. 11e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVQPG 8
DB 35 GRVLVQPG 42
1:|||||

RESULT 7
O9JLM1 PRELIMINARY; PRT; 104 AA.
AC O9JLM1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE TELOMERASE REVERSE TRANSCRIPTASE (FRAGMENT).
GN TERT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RA Hatakeyama S., Ishikawa F.;
RT "Identification of the mouse telomerase reverse transcriptase (mTERT)
RT promoter.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF157502; AAF42984.1; -
KW RNA-directed DNA polymerase.
FT NON_TER 104
SQ SEQUENCE 104 AA; 11998 MW; 4B649E63476D3D44 CRC64;

Query Match 82.9%; Score 34; DB 11; Length 104;
Best Local Similarity 87.5%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVQPG 8
DB 35 GRVLVQPG 42
1:|||||

RESULT 8
O56048 PRELIMINARY; PRT; 227 AA.
AC O56048;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-1998 (TREMBLrel. 06, Last annotation update)
DE CAPSID PROTEIN (FRAGMENT).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage.
CC NCBI_TaxID=12461;
RN NCBI_TaxID=12461;
RP SEQUENCE FROM N.A.
RC STRAIN=TK15/92/NEPAL.
RA Gouvea V., Snellings N., Cohen S.J., Warren R.L., Myint K.S.A.,
RA Shouva V., Vaughan D.W., Hoke C.H. Jr., Innis B.L.;
RL Virus Res. 0:0-0(1997).
DR EMBL: AF020604; AAB93555.1; -
FT NON_TER 1
FT NON_TER 227
SQ SEQUENCE 227 AA; 24525 MW; 887B492F9559F128 CRC64;

Query Match 82.9%; Score 34; DB 14; Length 227;
Best Local Similarity 85.7%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVLVQPG 8
DB 52 RVLVQPG 58
1:|||||

```

RESULT 9
ID 056049 PRELIMINARY; PRT; 227 AA.
AC 056049;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE 01-JUN-1998 (TREMBLrel. 06, Last annotation update)
DE CAPSID PROTEIN (FRAGMENT).
OS Hepatitis E virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TK5/95/NEPAL;
RA Gourea V., Snellings N., Cohen S.J., Warren R.L., Myint K.S.A.,
RA Shrestha M.P., Vaughn D.W., Hoke C.H. Jr., Innis B.L.;
RL Virus Res. 0:0-0(1997).
DR EMBL: AF020605; AAB93556.1; -.
FT NON_TER 1 227
SQ SEQUENCE 227 AA; 24585 MW; 3FAD6B04E634F65F CRC64;

```

```

Query Match
Best Local Similarity 82.9%; Score 34; DB 14; Length 227;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 RVLVOPG 8
Db 52 RILVOPG 58

```

```

RESULT 10
ID 09W827 PRELIMINARY; PRT; 227 AA.
AC 09W827;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE CASID PROTEIN (FRAGMENT).
OS Hepatitis E virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TK78/87/NEPAL, TK104/91/NEPAL, TK4/95/NEPAL, NEP4/94/NEPAL;
RA Gourea V., Snellings N., Cohen S.J., Warren R.L., Myint K.S.A.,
RA Shrestha M.P., Vaughn D.W., Hoke C.H. Jr., Innis B.L.;
RT "Hepatitis E virus in Nepal: similarities with the Burmese and Indian
variants.";
RL Virus Res. 0:0-0(1997).
DR EMBL: AF020608; AAB93559.1; -.
DR EMBL: AF020603; AAB93554.1; -.
DR EMBL: AF020606; AAB93557.1; -.
DR EMBL: AF020607; AAB93558.1; -.
FT NON_TER 1 227
SQ SEQUENCE 227 AA; 24555 MW; BFB61995A78A997 CRC64;

```

```

Query Match
Best Local Similarity 82.9%; Score 34; DB 14; Length 227;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 RVLVOPG 8
Db 52 RILVOPG 58

```

```

RESULT 11
O9PEW8

```

```

ID 09PEW8 PRELIMINARY; PRT; 253 AA.
AC 09PEW8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE UNQUINOL CYTOCHROME C OXIDOREDUCTASE, CYTOCHROME C1 SUBUNIT.
GN XFO910.
OS Xylella fastidiosa.
OC Bacteria: Proteobacteria; gamma subdivision; Xanthomonas group;
OX Xylella.
OC NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares R., Alves L.M.C., Arya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Gartner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA de Souza A.P., Terenzi M.F., Siqueira W.J., de Souza A.A.,
RA Vallada H., Van Sluys W.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Melandris J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-157(2000).
DR EMBL: AE003930; AAF83720.1; -.
DR EMBL: AE003930; CYTOCHROME_C1;
DR INTERPRO: IPR000345; -.
DR INTERPRO: IPR002326; -.
DR PFAM: PF02167; Cytochrome_C1; 2.
DR PRINTS: PR00603; CYTOCHROME_C1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 253 AA; 28431 MW; 78C4F596EADCDBA7 CRC64;

```

```

Query Match
Best Local Similarity 82.9%; Score 34; DB 2; Length 253;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 GRVLVOPG 8
Db 187 GLVLVOPG 194

```

```

RESULT 12
ID 09YB1 PRELIMINARY; PRT; 283 AA.
AC 09YB1;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE STRUCTURAL PROTEIN 2 (FRAGMENT).
OS Hepatitis E virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage.
OX NCBI_TaxID=12461;
RN [1]

```


RP SEQUENCE FROM N.A.
 RC STRAIN-2B NIH ABBOTTABAD, PAKISTAN;
 RA van Cuyck-Gandre H., Clements N.J., Zhang H.Y., Gaudill J.C.,
 RA Cohen S.G., Courssaget P., Buisson Y., Warren R.L., Longer C.F.;
 RT "Partial Sequence of HEV Isolates from North Africa and Pakistan:
 RT Comparison with Known HEV Sequences";
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U40045; AAD09446.1; -.
 FT NON_TER 1
 FT SEQUENCE 283 AA; 30345 MW; 921C296D7E141C33 CRC64;

Query Match 82.9%; Score 34; DB 14; Length 283;
 Best Local Similarity 85.7%; Pred. No. 42;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVLVOPG 8
 DB 100 RILVOPG 106

RESULT 13
 O9WI48 PRELIMINARY: PRT: 436 AA.
 AC O9WI48:
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE STRUCTURAL PROTEIN (FRAGMENT).
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage.
 OX NCBI_TaxID-12461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CHINA;
 RA Jiang L., Yang G., Shen X.L.;
 RT "Cloning and sequence analysis of the 1.3kb cDNA fragment from open
 RT reading frame 2 (ORF2) of hepatitis E virus (HEV).";
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF11552; AAD33878.1; -.
 FT NON_TER 1
 FT SEQUENCE 436 AA; 47256 MW; EDOAF486B072E24 CRC64;

Query Match 82.9%; Score 34; DB 14; Length 436;
 Best Local Similarity 85.7%; Pred. No. 66;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVLVOPG 8
 DB 9 RILVOPG 15

RESULT 14
 O9WNN1 PRELIMINARY: PRT: 466 AA.
 AC O9WNN1:
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE ORF2 (FRAGMENT).
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage.
 OX NCBI_TaxID-12461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CH-T21;
 RA Yang J., Zhang H., Wang Y., Li H.;
 RT "Identification of a Novel Hepatitis E virus Responsible for Sporadic
 RT Acute Hepatitis in China.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF151963; AAD41519.1; -.
 FT NON_TER 1
 FT SEQUENCE 466 AA; 50425 MW; 17695C0DABA63DDE CRC64;

Query Match 82.9%; Score 34; DB 14; Length 466;
 Best Local Similarity 85.7%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVLVOPG 8
 DB 40 RILVOPG 46

RESULT 15
 O9WNN2 PRELIMINARY: PRT: 486 AA.
 AC O9WNN2:
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE ORF2 (FRAGMENT).
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage.
 OX NCBI_TaxID-12461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CH-T11;
 RA Yang J., Zhang H., Wang Y., Li H.;
 RT "Identification of a Novel Hepatitis E virus Responsible for Sporadic
 RT Acute Hepatitis in China.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF151962; AAD41518.1; -.
 FT NON_TER 1
 FT SEQUENCE 486 AA; 52781 MW; 82B4D88570BFE61 CRC64;

Query Match 82.9%; Score 34; DB 14; Length 486;
 Best Local Similarity 85.7%; Pred. No. 74;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVLVOPG 8
 DB 59 RILVOPG 65

Search completed: June 13, 2001, 14:20:24
 Job time: 724 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 13, 2001, 14:16:31 ; Search time 118.55 Seconds
(without alignments)
1.296 Million cell updates/sec

Title: PCT-US01-05825a-3

Perfect score: 41

Sequence: 1 GRVLVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	90.2	118	2	US-08-652-816A-12
2	34	82.9	20	2	US-08-859-931A-4
3	34	82.9	246	1	US-07-843-125-11
4	34	82.9	435	1	US-08-259-148A-18
5	34	82.9	435	1	US-08-484-054-18
6	34	82.9	435	2	US-07-876-941A-18
7	34	82.9	436	1	US-08-259-148A-17
8	34	82.9	436	1	US-08-484-054-17
9	34	82.9	436	2	US-07-876-941A-17
10	34	82.9	525	5	PCT-US95-13703-27
11	34	82.9	525	5	PCT-US95-13703-28
12	34	82.9	540	5	PCT-US95-13703-25
13	34	82.9	540	5	PCT-US95-13703-26
14	34	82.9	549	5	PCT-US95-13703-15
15	34	82.9	549	5	PCT-US95-13703-16
16	34	82.9	659	1	US-08-240-049B-16
17	34	82.9	659	1	US-08-259-148A-16
18	34	82.9	659	1	US-08-484-054-20
19	34	82.9	659	2	US-07-876-941A-20
20	34	82.9	660	1	US-08-240-049B-15
21	34	82.9	660	1	US-08-259-148A-19
22	34	82.9	660	1	US-08-484-054-19
23	34	82.9	660	2	US-07-876-941A-19
24	34	82.9	660	2	US-08-840-316-2
25	34	82.9	660	4	US-08-478-507-8
26	34	82.9	660	4	US-08-809-523-2
27	34	82.9	660	5	PCT-US93-08849A-2

ALIGNMENTS

28	34	82.9	660	5	PCT-US93-08849-2	Sequence 2, Appl
29	34	82.9	660	5	PCT-US95-13703-13	Sequence 13, Appl
30	34	82.9	660	5	PCT-US95-13703-14	Sequence 14, Appl
31	31	75.6	115	1	US-08-468-661-1	Sequence 1, Appl
32	31	75.6	115	1	US-08-466-272A-1	Sequence 1, Appl
33	31	75.6	115	1	US-08-478-857-1	Sequence 1, Appl
34	31	75.6	115	2	US-08-471-771-1	Sequence 1, Appl
35	31	75.6	115	3	US-09-130-783-1	Sequence 1, Appl
36	31	75.6	118	3	US-08-545-809A-125	Sequence 125, App
37	30	73.2	118	3	US-08-545-809A-108	Sequence 108, App
38	30	73.2	119	2	US-08-475-000-16	Sequence 16, Appl
39	30	73.2	119	2	US-08-483-199-16	Sequence 16, Appl
40	30	73.2	119	2	US-08-484-508-16	Sequence 16, Appl
41	30	73.2	123	2	US-08-428-197-38	Sequence 38, Appl
42	30	73.2	123	5	PCT-US93-10555-38	Sequence 38, Appl
43	30	73.2	451	3	US-08-679-635A-2	Sequence 2, Appl
44	29	70.7	116	3	US-08-545-809A-143	Sequence 143, App
45	29	70.7	125	1	US-08-276-852-72	Sequence 72, Appl

RESULT 1
US-08-652-816A-12
Sequence 12, Application US/08652816A
Patent No. 5872215
GENERAL INFORMATION:
APPLICANT: Osbourn, JK
APPLICANT: Allen, DJ
TITLE OF INVENTION: Specific binding members, materials and
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESSES:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,816A
FILING DATE: 23-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9125579.4
FILING DATE: 02-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9125579.8
FILING DATE: 02-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206372.6
FILING DATE: 23-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9525004.9
FILING DATE: 07-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9610824.6
FILING DATE: 23-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/02240
FILING DATE: 02-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/244,597
FILING DATE: 01-JUN-1994
ATTORNEY/AGENT INFORMATION:

NAME: David W. Clough
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/33308
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 118 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-652-816A-12

Query Match 90.2%; Score 37; DB 2; Length 118;
Best Local Similarity 87.5%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GRVLVOPG 8
Db 8 GRVLVOPG 15

RESULT 2
US-08-859-931A-4
Sequence 4, Application US/08859931A
Patent No. 5945510

GENERAL INFORMATION:
APPLICANT: FASANO, Alessio
TITLE OF INVENTION: SUBSTANTIALLY PURE ZONULIN, A
TITLE OF INVENTION: PHYSIOLOGICAL MODULATOR OF
TITLE OF INVENTION: MAMMALIAN TIGHT JUNCTIONS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MIOW, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
CITY: Washington, D.C.
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,931A
FILING DATE: 21 May 1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: A-6901
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-859-931A-4

Query Match 82.9%; Score 34; DB 2; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.88;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GRVLVOPG 8
Db 8 GRVLVOPG 15

RESULT 3
US-07-843-125-11
Sequence 11, Application US/07843125
Patent No. 5395750

GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J
TITLE OF INVENTION: Methods for Producing Proteins which
TITLE OF INVENTION: Bind to Predetermined Antigens
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110-1199

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/843,125
FILING DATE: 19920228
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Roseman, Catherine R
REGISTRATION NUMBER: 34240
REFERENCE/DOCKET NUMBER: 8594
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-6208
TELEFAX: (201) 235-3500
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 246 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-843-125-11

Query Match 82.9%; Score 34; DB 1; Length 246;
Best Local Similarity 87.5%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GRVLVOPG 8
Db 8 GRVLVOPG 15

RESULT 4
US-08-259-148A-18
Sequence 18, Application US/08259148A
Patent No. 5741490

GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R.
APPLICANT: Bradley, Daniel W.
APPLICANT: Twu, Ji-Shin
APPLICANT: Purdy, Michael A.
APPLICANT: Krawczynski, Krzysztof Z.
APPLICANT: Yarbough, Patricia D.
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

Query Match 82.9%; Score 34; DB 1; Length 246;
Best Local Similarity 87.5%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/259,148A
FILING DATE: 13-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0860
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: C2, MEXICO, FIGURE 9
US-08-259-148A-18

Query Match 82.9%; Score 34; DB 1; Length 435;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLVOPG 8
Db 9 RLVOPG 15

RESULT 5
US-08-484-054-18
Sequence 18, Application US/08484054
Patent No. 5770689
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R.
APPLICANT: Bradley, Daniel W.
APPLICANT: Tuu, Jr-Shun
APPLICANT: Purdy, Michael A.
APPLICANT: Tam, Albert W.
APPLICANT: Krawczynski, Krzysztof Z.
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA

COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,054
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 870,985
FILING DATE: 20-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.38
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0860
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: C2, MEXICO, FIGURE 9
US-08-484-054-18

Query Match 82.9%; Score 34; DB 1; Length 435;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLVOPG 8
Db 9 RLVOPG 15

RESULT 6
US-07-876-941A-18
Sequence 18, Application US/07876941A
Patent No. 5885768
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R.
APPLICANT: Bradley, Daniel W.
APPLICANT: Mitchell, Carl
APPLICANT: Tam, Albert W.
TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and Antibodies
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/876,941A
FILING DATE: 01-MAY-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.33
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: C2, MEXICO, FIGURE 9
US-07-876-941A-18

Query Match 82.9%; Score 34; DB 2; Length 435;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVLVOPG 8
Db 9 RILVOPG 15

RESULT 7
US-08-259-148A-17
Sequence 17, Application US/08259148A
Patent No. 5741490
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R.
APPLICANT: Bradley, Daniel W.
APPLICANT: Twu, Jr-Shin
APPLICANT: Purdy, Michael A.
APPLICANT: Tam, Albert W.
APPLICANT: Krawczynski, Krzysztof Z.
APPLICANT: Varbough, Patrice D.

TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/259,148A
FILING DATE: 13-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 436 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 9
US-08-259-148A-17

Query Match 82.9%; Score 34; DB 1; Length 436;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVLVOPG 8
Db 9 RILVOPG 15

RESULT 8
US-08-484-054-17
Sequence 17, Application US/08484054
Patent No. 5770689
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R.
APPLICANT: Bradley, Daniel W.
APPLICANT: Twu, Jr-Shin
APPLICANT: Purdy, Michael A.

APPLICANT: Tam, Albert W.
APPLICANT: Krawczynski, Krzysztof Z.
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,054
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 870,985
FILING DATE: 20-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.38
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 436 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 9
US-08-484-054-17

Query Match 82.9%; Score 34; DB 1; Length 436;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVLVOPG 8
1:|||||
DB 9 RVLVOPG 15

RESULT 9
US-07-876-941A-17
; Sequence 17, Application US/07876941A
; Patent No. 5885768

GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R.
APPLICANT: Bradley, Daniel W.
APPLICANT: Tam, Albert W.
APPLICANT: Mitchell, Carl
TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
TITLE OF INVENTION: Antibodies
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/876,941A
FILING DATE: 01-MAY-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.33
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 436 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 9
US-07-876-941A-17

Query Match 82.9%; Score 34; DB 2; Length 436;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVLVOPG 8
1:|||||
DB 9 RVLVOPG 15

RESULT 10
PCT-US95-13703-27
; Sequence 27, Application PC/7US9513703

GENERAL INFORMATION:

APPLICANT: HEPATITIS E VIRUS ANTIGENS AND
TITLE OF INVENTION: USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13703
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.41
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 525 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)
PCT-US95-13703-27

Query Match 82.9%; Score 34; DB 5; Length 525;
Best Local Similarity 85.7%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 RVLVQPG 8
DB 122 RILVQPG 128

RESULT 11
PCT-US95-13703-28
Sequence 28, Application PC/TUS9513703
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13703

FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.41
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 525 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Hepatitis E virus (Mexico strain)
PCT-US95-13703-28

Query Match 82.9%; Score 34; DB 5; Length 525;
Best Local Similarity 85.7%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 RVLVQPG 8
DB 122 RILVQPG 128

RESULT 12
PCT-US95-13703-25
Sequence 25, Application PC/TUS9513703
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13703
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.41
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)

INDIVIDUAL ISOLATE: r62kda, 58.1 kda
PCT-US95-13703-25

Query Match 82.9%; Score 34; DB 5; Length 540;
Best Local Similarity 85.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVLVOPG 8
DB 122 RILVOPG 128

RESULT 13
PCT-US95-13703-26
; Sequence 26, Application PC/TUS9513703
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
; TITLE OF INVENTION: USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13703
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E virus (Mexico strain)
; INDIVIDUAL ISOLATE: r62kda, 58.1 kda
PCT-US95-13703-26

Query Match 82.9%; Score 34; DB 5; Length 540;
Best Local Similarity 85.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVLVOPG 8
DB 122 RILVOPG 128

RESULT 14
PCT-US95-13703-15
; Sequence 15, Application PC/TUS9513703
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND

TITLE OF INVENTION: USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13703
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.41
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 549 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)
INDIVIDUAL ISOLATE: r62kda, FIGURE 4
PCT-US95-13703-15

Query Match 82.9%; Score 34; DB 5; Length 549;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVLVOPG 8
DB 122 RILVOPG 128

RESULT 15
PCT-US95-13703-16
; Sequence 16, Application PC/TUS9513703
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
; TITLE OF INVENTION: USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13703
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.41
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 549 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Hepatitis E virus (Mexico strain)
INDIVIDUAL ISOLATE: r62kDa, FIGURE 4
PCT-US95-13703-16

Query Match 82.9%; Score 34; DB 5; Length 549;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 RLVOPG 8
1:|||||
Db 122 RLVOPG 128

Search completed: June 13, 2001, 14:16:32
Job time: 493 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:14:31 ; Search time 229.28 Seconds

(without alignments)
1.995 Million cell updates/sec

Title: PCT-US01-05825A-4

Perfect score: 40

Sequence: 1 GRVLVQDG 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SIDS6/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	8	21	Y79108
2	37	92.5	8	21	Y79112
3	35	87.5	8	21	Y79106
4	34	85.0	442	21	G23560
5	34	85.0	493	21	G23559
6	34	85.0	495	21	G23558
7	34	85.0	495	21	G51350
8	34	85.0	528	21	G51349
9	34	85.0	548	21	G51348
10	33	82.5	8	21	Y79107
11	33	82.5	8	21	Y79116

ALIGNMENTS

12	33	82.5	8	21	Y79120	Peptide antagonist
13	33	82.5	137	21	G05180	Arabidopsis thalia
14	33	82.5	149	21	G36079	Zea mays protein f
15	33	82.5	151	21	G44702	Zea mays protein f
16	33	82.5	151	21	G23109	Arabidopsis thalia
17	33	82.5	166	21	B28135	Meristem-specific
18	33	82.5	166	21	G07728	Arabidopsis thalia
19	33	82.5	166	21	G11107	Arabidopsis thalia
20	33	82.5	166	21	G20486	Arabidopsis thalia
21	33	82.5	166	21	G23108	Arabidopsis thalia
22	33	82.5	166	21	G45133	Arabidopsis thalia
23	33	82.5	166	21	G45692	Arabidopsis thalia
24	33	82.5	166	21	G45717	Arabidopsis thalia
25	33	82.5	168	21	G36078	Zea mays protein f
26	33	82.5	168	21	G44701	Zea mays protein f
27	33	82.5	180	21	G45132	Arabidopsis thalia
28	32	80.0	8	21	Y79110	Peptide antagonist
29	32	80.0	399	13	R20006	Zonula occludens t
30	32	80.0	415	21	G59077	Arabidopsis thalia
31	32	80.0	415	21	G59619	Arabidopsis thalia
32	32	80.0	449	21	G59076	Arabidopsis thalia
33	32	80.0	463	21	G59618	Arabidopsis thalia
34	32	80.0	575	21	G59617	Arabidopsis thalia
35	32	80.0	590	21	G59075	Arabidopsis thalia
36	31	77.5	117	16	R66214	Protein encoded by
37	31	77.5	117	17	R81470	Nocardia corallina
38	31	77.5	144	22	B19779	Plomycus equi car
39	31	77.5	305	22	B19781	Plomycus equi car
40	31	77.5	478	22	B19778	Plomycus equi car
41	31	77.5	572	19	W68488	Mouse ULRP-4 prote
42	30	75.0	8	21	Y79111	Peptide antagonist
43	30	75.0	8	21	Y79124	Peptide antagonist
44	30	75.0	79	20	Y33805	Scsmb-YERO29 a yea
45	30	75.0	96	21	G22719	Zea mays protein f

RESULT 1
Y79108 standard; Peptide; 8 AA.
Y79108:
AC Y79108:
DT 05-JUN-2000 (first entry)
XX
DE Peptide antagonist of zonulin.
XX
KW Zonulin; antagonist; zonula occludens toxin receptor;
KW blood-brain barrier; antiinflammatory; cerebroprotective;
KW neuroprotective; dermatological; antitumor; antiviral;
KW antibacterial; cytostatic; anti-HIV; vulnery; anti allergic;
KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW gastrointestinal inflammation; therapy.
XX
OS Synthetic.
XX
FN WO200007609-A1.
XX
PD 17-FEB-2000.
XX
XX 28-JUL-1999; 99WO-US16683.
XX
PR 03-AUG-1998; 98US-0127815.
XX
PA (UYMA-) UNITV MARYLAND BALTIMORE.
XX
PI Fasano A;
XX
XX WPI; 2000-205565/18.
DR
XX
PT New peptide antagonist of zonulin useful as antiinflammatory agent for

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS
 PS Claim 1: Page 42: 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infection, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 CC Sequence 8 AA:
 SQ

Query Match 100.0%; Score 40; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVLVODG 8
 |||||
 Db 1 grvlvqdg 8

RESULT 2
 ID Y79112 standard; Peptide: 8 AA.
 AC Y79112;
 XX

DT 05-JUN-2000 (first entry)
 XX

DE Peptide antagonist of zonulin.
 XX

KW Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiulcer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.
 XX

OS Synthetic.
 XX

PN WO200007609-A1.
 XX

PD 17-FEB-2000.
 XX

PF 28-JUL-1999; 99WO-US16683.
 XX

PR 03-AUG-1998; 98US-0127815.
 XX

PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX

PI Fasano A;
 XX
 DR WPI; 2000-205565/18.
 XX

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS
 PS Claim 1: Page 43: 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infection, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 CC Sequence 8 AA:
 SQ

Query Match 92.5%; Score 37; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVLVODG 8
 |||||
 Db 1 grvlvqdg 8

RESULT 3
 ID Y79106 standard; Peptide: 8 AA.
 AC Y79106;
 XX

DT 05-JUN-2000 (first entry)
 XX

DE Peptide antagonist of zonulin.
 XX

KW Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiulcer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.
 XX

OS Synthetic.
 XX

PN WO200007609-A1.
 XX

PD 17-FEB-2000.
 XX

PF 28-JUL-1999; 99WO-US16683.
 XX

XX	03-AUG-1998:	9805-0127815.
PR		
XX	(UYMA-).	UNIV MARYLAND BALTIMORE.
PA		
XX		
PI	Fasano A:	
DR	WPI: 2000-205565/18.	
XX		
PT	New peptide antagonist of zonulin useful as antiinflammatory agent for	
PT	treating cerebral ischemia, stroke, cerebral edema, gastritis,	
PT	shigellosis, viral gastroenteritis, meningitis, encephalomyelitis	
PS	Claim 1: Page 41: 69pp: English.	
XX		
CC	This present sequence is that of a peptide antagonist of zonulin	
CC	(2), one of 25 such peptides (see Y79105-23) of the invention,	
CC	which bind to a zonula occludens toxin (ZOT) receptor, yet do not	
CC	physiologically modulate the opening of mammalian tight junctions	
CC	(77). The peptide antagonists are based on a common motif of ZOT	
CC	and human zonulins, which is believed to be critical for receptor	
CC	binding. They can be prepared by chemical synthesis or by use of	
CC	recombinant DNA techniques. The peptide antagonists are used as an	
CC	antiinflammatory agents in the treatment of gastrointestinal	
CC	inflammation, where they bind to the ZOT receptor in the intestine	
CC	and yet does not physiologically modulate the opening of TJ in the	
CC	intestine. Gastrointestinal inflammation conditions give rise to	
CC	increased intestinal permeability and the peptide is useful for	
CC	treating intestinal conditions that cause protein losing enteropathy	
CC	caused by infection, e.g. Clostridium difficile infection,	
CC	enterocolitis, shigellosis, viral gastroenteritis, parasite	
CC	infestation, bacterial overgrowth, whipple's disease, diseases with	
CC	mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,	
CC	cologenous colitis, inflammatory bowel disease, diseases marked by	
CC	lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,	
CC	sarcoïdosis lymphoma, mesenteric tuberculosis, and after surgical	
CC	correction of congenital heart disease with Fontan's operation,	
CC	mucosal diseases without ulceration, e.g. Menetrier's disease,	
CC	colic disease, eosinophilic gastroenteritis, and immune diseases,	
CC	e.g. systemic lupus erythematosus or food allergies, primarily to	
CC	milk.	
XX		
XX	Sequence 8 AA:	
SO		
XX		
QY	1 GRVLVQDG 8	87.5%: Score 35; DB 21; Length 8;
		Best Local Similarity 87.5%: Pred. NO. 3.2e+05;
Db	1 grvcvqdg 8	Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
RESULT 4		
ID	G23560	
XX	G23560 standard; protein; 442 AA.	
AC	G23560:	
XX		
DT	17-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana protein fragment SPO ID NO: 26913.	
XX		
XX	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
XX	termination sequence.	
OS	Arabidopsis thaliana.	
XX	EP1033405-A2.	
PN		
XX		
DP	06-SEP-2000.	

XX	25-FEB-2000	2000EP-0301439
XX	25-FEB-1999	9905-0121885
PR	05-MAR-1999	9905-0123180
PR	09-MAR-1999	9905-0123348
PR	23-MAR-1999	9905-0125788
PR	25-MAR-1999	9905-0126564
PR	29-MAR-1999	9905-0126785
PR	01-APR-1999	9905-0127462
PR	06-APR-1999	9905-0128224
PR	08-APR-1999	9905-0128714
PR	16-APR-1999	9905-0129845
PR	19-APR-1999	9905-0130077
PR	21-APR-1999	9905-0130449
PR	23-APR-1999	9905-0130510
PR	23-APR-1999	9905-0130691
PR	28-APR-1999	9905-0131449
PR	30-APR-1999	9905-0132048
PR	30-APR-1999	9905-0132407
PR	04-MAY-1999	9905-0132484
PR	05-MAY-1999	9905-0132485
PR	05-MAY-1999	9905-0132486
PR	07-MAY-1999	9905-0132487
PR	07-MAY-1999	9905-0132683
PR	11-MAY-1999	9905-0134226
PR	11-MAY-1999	9905-0134218
PR	14-MAY-1999	9905-0134218
PR	14-MAY-1999	9905-0134221
PR	14-MAY-1999	9905-0134370
PR	16-MAY-1999	9905-0134768
PR	19-MAY-1999	9905-0134941
PR	19-MAY-1999	9905-0135124
PR	20-MAY-1999	9905-0135124
PR	24-MAY-1999	9905-0135253
PR	25-MAY-1999	9905-0135629
PR	27-MAY-1999	9905-0136021
PR	28-MAY-1999	9905-0136392
PR	01-JUN-1999	9905-0137282
PR	03-JUN-1999	9905-0137622
PR	04-JUN-1999	9905-0137502
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138054
PR	10-JUN-1999	9905-0138400
PR	10-JUN-1999	9905-0138477
PR	14-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-0139452
PR	17-JUN-1999	9905-0139453
PR	18-JUN-1999	9905-0139492
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139461
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139463
PR	18-JUN-1999	9905-0139750
PR	18-JUN-1999	9905-0139763
PR	21-JUN-1999	9905-0139617
PR	22-JUN-1999	9905-0139689
PR	23-JUN-1999	9905-0140353
PR	23-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140655
PR	28-JUN-1999	9905-0140823
PR	29-JUN-1999	9905-0140921
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141854
PR	01-JUL-1999	9905-0142184
PR	02-JUL-1999	9905-0142055
PR	06-JUL-1999	9905-0142390

```

PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154018.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.

```

```

PR 24-SEP-1999; 99US-0155559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158332.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      85.0%; Score 34; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 RVLVDG 8
Db 250 rvlvgdg 256

```

```

RESULT 5
G23559 ID G23559 standard; Protein; 493 AA.
AC G23559;
XX
XX 17-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 26912.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.

```

PR	25-MAR-1999;	9905-0126264;	PR	16-JUL-1999;	9905-0144086;
PR	25-MAR-1999;	9905-0126785;	PR	16-JUL-1999;	9905-0144325;
PR	01-APR-1999;	9905-0127462;	PR	19-JUL-1999;	9905-0144331;
PR	06-APR-1999;	9905-0128234;	PR	19-JUL-1999;	9905-0144332;
PR	08-APR-1999;	9905-0128714;	PR	19-JUL-1999;	9905-0144333;
PR	16-APR-1999;	9905-0129845;	PR	19-JUL-1999;	9905-0144334;
PR	19-APR-1999;	9905-0130077;	PR	19-JUL-1999;	9905-0144335;
PR	21-APR-1999;	9905-0130049;	PR	20-JUL-1999;	9905-0144332;
PR	23-APR-1999;	9905-0130510;	PR	20-JUL-1999;	9905-0144632;
PR	23-APR-1999;	9905-0130891;	PR	20-JUL-1999;	9905-0144684;
PR	28-APR-1999;	9905-0131449;	PR	21-JUL-1999;	9905-0144814;
PR	30-APR-1999;	9905-0132048;	PR	21-JUL-1999;	9905-0144816;
PR	30-APR-1999;	9905-0132407;	PR	21-JUL-1999;	9905-0145088;
PR	04-MAY-1999;	9905-0132484;	PR	21-JUL-1999;	9905-0145088;
PR	05-MAY-1999;	9905-0132485;	PR	22-JUL-1999;	9905-0145087;
PR	06-MAY-1999;	9905-0132486;	PR	22-JUL-1999;	9905-0145087;
PR	06-MAY-1999;	9905-0132487;	PR	22-JUL-1999;	9905-0145089;
PR	07-MAY-1999;	9905-0132863;	PR	22-JUL-1999;	9905-0145192;
PR	11-MAY-1999;	9905-0134256;	PR	23-JUL-1999;	9905-0145145;
PR	14-MAY-1999;	9905-0134218;	PR	23-JUL-1999;	9905-0145218;
PR	14-MAY-1999;	9905-0134219;	PR	23-JUL-1999;	9905-0145224;
PR	14-MAY-1999;	9905-0134221;	PR	26-JUL-1999;	9905-0145276;
PR	14-MAY-1999;	9905-0134370;	PR	27-JUL-1999;	9905-0145913;
PR	18-MAY-1999;	9905-0134768;	PR	27-JUL-1999;	9905-0145918;
PR	19-MAY-1999;	9905-0134941;	PR	27-JUL-1999;	9905-0145919;
PR	20-MAY-1999;	9905-0135124;	PR	28-JUL-1999;	9905-0145951;
PR	21-MAY-1999;	9905-0135353;	PR	02-AUG-1999;	9905-0146386;
PR	24-MAY-1999;	9905-0135629;	PR	02-AUG-1999;	9905-0146389;
PR	25-MAY-1999;	9905-0136021;	PR	02-AUG-1999;	9905-0146388;
PR	27-MAY-1999;	9905-0136392;	PR	03-AUG-1999;	9905-0147038;
PR	28-MAY-1999;	9905-0136782;	PR	04-AUG-1999;	9905-0147204;
PR	01-JUN-1999;	9905-0137222;	PR	04-AUG-1999;	9905-0147303;
PR	03-JUN-1999;	9905-0137528;	PR	05-AUG-1999;	9905-0147192;
PR	04-JUN-1999;	9905-0137502;	PR	05-AUG-1999;	9905-0147226;
PR	07-JUN-1999;	9905-0137724;	PR	06-AUG-1999;	9905-0147303;
PR	08-JUN-1999;	9905-0138094;	PR	06-AUG-1999;	9905-0147416;
PR	10-JUN-1999;	9905-0138540;	PR	09-AUG-1999;	9905-0147493;
PR	10-JUN-1999;	9905-0138847;	PR	09-AUG-1999;	9905-0147935;
PR	14-JUN-1999;	9905-0139119;	PR	10-AUG-1999;	9905-0148171;
PR	16-JUN-1999;	9905-0139452;	PR	11-AUG-1999;	9905-0148341;
PR	17-JUN-1999;	9905-0139453;	PR	12-AUG-1999;	9905-0148341;
PR	17-JUN-1999;	9905-0139457;	PR	13-AUG-1999;	9905-0148565;
PR	18-JUN-1999;	9905-0139459;	PR	13-AUG-1999;	9905-0148684;
PR	18-JUN-1999;	9905-0139459;	PR	16-AUG-1999;	9905-0148368;
PR	18-JUN-1999;	9905-0139455;	PR	17-AUG-1999;	9905-0149175;
PR	18-JUN-1999;	9905-0139456;	PR	18-AUG-1999;	9905-0149426;
PR	18-JUN-1999;	9905-0139457;	PR	20-AUG-1999;	9905-0149722;
PR	18-JUN-1999;	9905-0139458;	PR	20-AUG-1999;	9905-0149723;
PR	18-JUN-1999;	9905-0139459;	PR	20-AUG-1999;	9905-0149929;
PR	18-JUN-1999;	9905-0139460;	PR	23-AUG-1999;	9905-0149902;
PR	18-JUN-1999;	9905-0139461;	PR	23-AUG-1999;	9905-0149930;
PR	18-JUN-1999;	9905-0139462;	PR	25-AUG-1999;	9905-0150566;
PR	18-JUN-1999;	990			

PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

Best Local Similarity 85.0%; Score 34; DB 21; Length 493;

Matches 7; Conservative 0; Mismatches 0;

Indels 0; Caps 0;

OY 2 RVLVQDG 8
|||||
DB 301 rvlvqdg 307

RESULT 6

G23558 standard; Protein; 495 AA.

AC G23558;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 26911.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134769.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154773.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.0%; Score 34; DB 21; Length 495;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVLVDPG 8
Db 303 rvlvdpdg 309
IIIIII

RESULT 7
G51350
ID G51350 standard; Protein; 495 AA.
XX
AC G51350;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65163.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 200DEP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130891.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 07-MAY-1999; 9905-0132487.
PR 11-MAY-1999; 9905-0132863.
PR 14-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139753.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139817.
PR 23-JUN-1999; 9905-0139817.
PR 23-JUN-1999; 9905-0140353.
PR 24-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140823.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141287.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142927.
PR 13-JUL-1999; 9905-0142977.
PR 14-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144085.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144632.
PR 21-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.

PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145218.
PR 26-JUL-1999; 9905-0145224.
PR 27-JUL-1999; 9905-0145226.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 28-JUL-1999; 9905-0145918.
PR 28-JUL-1999; 9905-0145918.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149902.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151066.
PR 30-AUG-1999; 9905-0151080.
PR 31-AUG-1999; 9905-0151303.
PR 01-SEP-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154719.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157173.
PR 06-OCT-1999; 9905-0157753.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159637.
PR 18-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.0%; Score 34; DB 21; Length 495;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 RVLVQDG 8
Db 303 rvlvqdg 309

RESULT 8
G51349
ID G51349 standard; Protein; 528 AA.
XX
AC G51349;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65162.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123348.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143547.
PR 14-JUL-1999; 99US-0143622.
PR 15-JUL-1999; 99US-0143605.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145118.
PR 23-JUL-1999; 99US-0145215.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.0%; Score 34; DB 21; Length 528;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVLVDDG 8
Db 336 rvlvddg 342

RESULT 9
G51348
ID G51348 standard; Protein; 548 AA.

AC G51348;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 65161.

KW Protein identification; signal transduction pathway; metabolic pathway;
KM hydridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 14-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 02-JUL-1999;	99US-0142154.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.

PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 05-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 06-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 09-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154479.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.
PR 26-OCT-1999;	99US-0161360.
PR 26-OCT-1999;	99US-0161361.
PR 28-OCT-1999;	99US-0161920.
PR 28-OCT-1999;	99US-0161922.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.0%; Score 34; DB 21; Length 548;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RVLVQDG 8
Db 356 rvlvqdg 362

RESULT 10

ID Y79107 standard; Peptide: 8 AA.

XX Y79107;

DT 05-JUN-2000 (first entry)

DE Peptide antagonist of zonulin.

XX Zonulin: antagonist; zonula occludens toxin receptor;
KM blood-brain barrier; antiinflammatory; cerebroprotective;
KM neuroprotective; dermatological; antiulcer; antiviral;
KM antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
KM hypotensive; immunosuppressive; antiparasitic; vasotropic;
KM gastrointestinal inflammation; therapy.

XX Synthetic.

OS WO200007609-A1.

PN 17-FEB-2000.

PD 28-JUL-1999; 99WO-US16683.

XX 03-AUG-1998; 98US-0127815.

PR (UYMA-) UNIV MARYLAND BALTIMORE.

PA Fasano A;

XX WPI; 2000-205565/18.

XX New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
PS Claim 1; Page 41; 69pp; English.

XX This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,

CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.
XX Sequence 8 AA.

Query Match 82.5%; Score 33; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. No. 32e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVQDG 8
Db 1 grvlvqdg 8

RESULT 11

ID Y79116 standard; Peptide: 8 AA.

XX Y79116;

DT 05-JUN-2000 (first entry)

DE Peptide antagonist of zonulin.

XX Zonulin: antagonist; zonula occludens toxin receptor;
KM blood-brain barrier; antiinflammatory; cerebroprotective;
KM neuroprotective; dermatological; antiulcer; antiviral;
KM antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
KM hypotensive; immunosuppressive; antiparasitic; vasotropic;
KM gastrointestinal inflammation; therapy.

XX Synthetic.

OS WO200007609-A1.

PN 17-FEB-2000.

PD 28-JUL-1999; 99WO-US16683.

XX 03-AUG-1998; 98US-0127815.

PR (UYMA-) UNIV MARYLAND BALTIMORE.

PA Fasano A;

XX WPI; 2000-205565/18.

XX New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
PS Claim 1; Page 44; 69pp; English.

XX This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with

CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer.
CC collagenous colitis, inflammatory bowel disease, diseases marked by,
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.
CC
SQ Sequence 8 AA:

Query Match 82.5%; Score 33; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. NO. 3.2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVQDG 8
Db 1 grglvqdg 8
1 ||||||
1 grglvqdg 8

RESULT 12
ID V79120 standard; Peptide; 8 AA.
XX V79120;
XX 05-JUN-2000 (first entry)
XX
XX Peptide antagonist of zonulin.
XX
XX Zonulin; antagonist; zonula occludens toxin receptor;
XX blood-brain barrier; antiinflammatory; cerebroprotective;
XX neuroprotective; dermatological; antidiarr; antiviral;
XX antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
XX hypotensive; immunosuppressive; antiparasitic; vasotropic;
XX gastrointestinal inflammation; therapy.
XX
XX Synthetic.
XX
XX WO20007609-A1.
XX
XX 17-FEB-2000.
XX
XX 28-JUL-1999; 99WO-US16683.
XX
XX 03-AUG-1998; 98US-0127815.
XX
XX (UYMA-) UNIV MARYLAND BALTIMORE.
XX
XX Fasano A;
XX
XX WPI: 2000-205565/18.
XX
XX
XX Claim 1; Page 46; 69pp; English.
XX
XX This present sequence is that of a peptide antagonist of zonulin
XX (2), one of 25 such peptides (see V79105-29) of the invention,
XX which bind to zonula occludens toxin (ZOT) receptor, yet do not
XX physiologically modulate the opening of mammalian tight junctions
XX (TJ). The peptide antagonists are based on a common motif of ZOT
XX and human zonulins, which is believed to be critical for receptor
XX binding. They can be prepared by chemical synthesis or by use of
XX recombinant DNA techniques. The peptide antagonists are used as an
XX antiinflammatory agents in the treatment of gastrointestinal
XX inflammation, where they bind to the ZOT receptor in the intestine
XX and yet does not physiologically modulate the opening of TJ in the
XX intestine. Gastrointestinal inflammation conditions give rise to

CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.
CC
SQ Sequence 8 AA:

Query Match 82.5%; Score 33; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. NO. 3.2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVQDG 8
Db 1 grglvqdg 8
1 ||||||
1 grglvqdg 8

RESULT 13
ID G05180 standard; Protein; 137 AA.
XX G05180;
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 1489.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hydridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137328.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 21-JUN-1999; 99US-0139750.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140981.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 82.5%; Score 33; DB 21; Length 137;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
Db 64 grvlvvdg 71

RESULT 14
G36079 ID G36079 standard; Protein; 149 AA.

AC G36079;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 44165.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148655.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149425.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 82.5%; Score 33; DB 21; Length 149;
Best Local Similarity 87.5%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GRVLVDG 8
Db 47 grvlvvg 54
RESULT 15
G44702 standard; Protein; 149 AA.
ID G44702;
XX G44702;
AC XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 56025.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
OS Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 82.5%; Score 33; DB 21; Length 149;
Best Local Similarity 87.5%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy	1	GRVLVODG	8
Db	47	grvlvvdg	54

Search completed: June 13, 2001, 14:14:32
Job time: 373 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:10:38 ; Search time 130.61 Seconds
(without alignments)
4.209 Million cell updates/sec

Title: PCT-US01-05825A-4

Perfect score: 40
Sequence: 1 GRVLVVDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	82.5	160	2 A75466	2-demethylmenaquin
2	33	82.5	161	2 S40872	2-demethylmenaquin
3	33	82.5	225	1 E53402	serine O-acetyltra
4	33	82.5	355	2 A26984	std protein - Str
5	33	82.5	482	2 H69392	4-hydroxybutyrate
6	33	82.5	593	2 I46526	sodium/dicarboxyla
7	33	82.5	893	2 T29621	hypothetical prote
8	33	82.5	162	2 G64153	2-demethylmenaquin
9	32	80.0	196	2 C64891	ferritinocellin-bin
10	32	80.0	217	1 B53402	serine O-acetyltra
11	32	80.0	310	2 C83267	hypothetical prote
12	32	80.0	320	2 I02398	hypothetical 36.2K
13	32	80.0	320	2 B44059	36K protein - appl
14	32	80.0	399	2 B82197	zona occludens tox
15	32	80.0	399	2 A43864	N-ethylmenelline ch
16	32	80.0	427	2 G69067	hypothetical prote
17	32	80.0	590	2 T06626	hypothetical prote
18	32	80.0	687	2 T29148	hypothetical prote
19	32	80.0	1302	2 T20767	hypothetical prote
20	32	80.0	1767	2 T20766	hypothetical prote
21	31	77.5	182	2 A64834	fibrital-like prot
22	31	77.5	234	2 S30998	gene 53 protein -
23	31	77.5	235	2 B72806	gpi3 protein - Myc
24	31	77.5	273	2 A72378	conserved hypotnet
25	31	77.5	309	2 E69200	conserved hypotnet
26	31	77.5	563	2 A69491	probable formylmet
27	31	77.5	575	2 T45668	hypothetical prote
28	31	77.5	582	2 B82724	cardiolipin syntha
29	31	77.5	656	2 B47096	hy1B homolog - Str

30	31	77.5	722	2 B75608	GMC oxidoreductase
31	31	77.5	823	2 T06574	Probable 1,4-alpha
32	31	77.5	1100	2 T30303	Ca2+-ATPase - Tryp
33	31	77.5	1166	2 T29009	hypothetical prote
34	31	77.5	1317	2 S77517	DNA-directed RNA p
35	31	77.5	1665	2 T29008	hypothetical prote
36	30	75.0	21	2 S33287	DNA polymerase - T
37	30	75.0	151	2 A83333	Probable deaminase
38	30	75.0	162	2 B83423	probable methyltra
39	30	75.0	162	2 F82084	Probable 2-demethy
40	30	75.0	180	2 S09888	hypothetical prote
41	30	75.0	193	2 D82271	hypothetical prote
42	30	75.0	196	2 S50487	yfnc protein VC086
43	30	75.0	213	2 T37039	hypothetical prote
44	30	75.0	218	2 T27701	hypothetical prote
45	30	75.0	231	1 A29459	orate phosphorib

ALIGNMENTS

RESULT 1
A75466
2-demethylmenaquinone 2-C-methyltransferase (EC 2.1.1.-) DR0859 [similarity] - Deinoc
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence-revision 03-Dec-1999 #text-change 02-Sep-2000
C:Accession: A75466
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: A75466
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-160 <WHI>
A:Cross-references: GB:AE001940; GB:AE000513; NID:g6458577; PIDN:AAF10437.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0859
A:Map position: 1
C:Keywords: methyltransferase

Query Match 82.5%; Score 33; DB 2; Length 160;
Best Local Similarity 87.5%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVLVVDG 8
DB 59 GRVLVVDG 66

RESULT 2
S40872
2-demethylmenaquinone 2-C-methyltransferase (EC 2.1.1.-) meng [validated] - Escherich
N:Alternate names: hypothetical protein fl61; menaquinone biosynthesis protein meng;
C:Species: Escherichia coli
C>Date: 19-May-1994 #sequence-revision 01-Sep-1995 #text-change 02-Sep-2000
R:Plunkett III, G.; Burland, V.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 21, 3391-3398, 1993
A>Title: Analysis of the Escherichia coli genome. III. DNA sequence of the region fro
A:Reference number: S40802; MUID:93347969
A:Accession: S40872
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-161 <PLU>
A:Cross-references: EMBL:L19201; NID:g304961; PIDN:AAB03061.1; PID:g305032
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1993
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: D65199

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-161 <BLAT>

A:Cross-references: GB:AE000467; GB:U00096; NID:q1790356; PIDN:AAC76911.1; PID:q1790364;

A:Experimental source: strain K-12, substrain MG1655

R:Hudepeth, M.E.S.; Suvana, K.; Meganathan, R.

submitted to the EMBL Data Library, April 1996

A:Description: Menquinone biosynthesis: identification of *E. coli* menA and meng.

A:Reference number: Z24499

A:Accession: T48648

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-161 <HND>

A:Cross-references: EMBL:U56082; PIDN:AA801208.1

A:Experimental source: strain K12

C:Genetics:

A:Gene: meng

A:Note: between glpK and hslU

C:Function:

A:Description: methyltransferase; catalyzes the conversion of demethyl-menaquinone to me

A:Pathway: menaquinone biosynthesis

C:Keywords: methyltransferase

Query Match 82.5%; Score 33; DB 2; Length 161;
Best Local Similarity 87.5%; Pred. No. 10;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
|||:|:|
DB 58 GRVLVODG 65

RESULT 3
E53402
serine O-acetyltransferase (EC 2.3.1.30) cyse [similarity] - *Bacillus stearothermophilus*

C:Species: *Bacillus stearothermophilus*

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-May-2000

C:Accession: E53402

R:Gagnon, Y.; Breton, R.; Puter, H.; Pelchat, M.; Grunberg-Manago, M.; Lapointe, J.

J. Biol. Chem. 269, 7473-7482, 1994

A:Title: Clustering and co-transcription of the *Bacillus subtilis* genes encoding the am

A:Reference number: A53402; MUID:94171772

A:Accession: E53402

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-225 <GAG>

C:Genetics:

A:Gene: cyse

C:Function:

A:Pathway: cysteine biosynthesis

A:Note: rate-limiting step

C:Superfamily: *Bacillus* serine acetyltransferase; serine acetyltransferase homology

C:Keywords: acyltransferase; coenzyme A

F:9-169/Domain: serine acetyltransferase homology <SAT>

Query Match 82.5%; Score 33; DB 1; Length 225;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
|||:|:|
DB 166 GRVLVODG 173

RESULT 4
A26984

strd protein - *Streptomyces griseus*

N:Alternate names: hypothetical hexose nucleotidylating enzyme

C:Species: *Streptomyces griseus*

C>Date: 20-Feb-1993 #sequence_revision 20-Feb-1993 #text_change 29-Sep-1999

C:Accession: A26984; A43701

R:Distler, J.; Ebert, A.; Mansouri, K.; Pissowolzh, K.; Stockmann, M.; Piepersberg,

Nucleic Acids Res. 15, 8041-8056, 1987

A:Title: Gene cluster for streptomycin biosynthesis in *Streptomyces griseus*: nucleoti

A:Reference number: A93676; MUID:88040426

A:Accession: A26984

A:Molecule type: DNA

A:Residues: 1-355 <DIS>

A:Cross-references: GB:Y00459

R:Distler, J.

submitted to the EMBL Data Library, December 1987

A:Reference number: A43701

A:Accession: A43701

A:Molecule type: DNA

A:Residues: 1-340 'N', 342-355 <DIS>

A:Cross-references: GB:Y00459; NID:q1621271; PIDN:CAA68514.1; PID:g49261

C:Superfamily: mannose-1-phosphate guanylyltransferase

Query Match 82.5%; Score 33; DB 2; Length 355;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
|||:|:|
DB 254 GRVLVODG 261

RESULT 5
H69392

4-hydroxybutyrate CoA transferase (cat-1) homolog - *Archaeoglobus fulgidus*

C:Species: *Archaeoglobus fulgidus*

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 17-Mar-2000

C:Accession: H69392

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod

Godok, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 350, 364-370, 1997

A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Attlich, P.; Kaine, B.P.; Sykes,

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch

A:Reference number: A69250; MUID:96049343

A:Accession: H69392

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-482 <KLE>

A:Cross-references: GB:AE001025; GB:AE000782; NID:q2689348; PIDN:AB90101.1; PID:g264

C:Superfamily: acetyl-CoA hydrolase

Query Match 82.5%; Score 33; DB 2; Length 482;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
|||:|:|
DB 225 GRVLVODG 232

RESULT 6
I46528
sodium/dicarboxylate cotransporter - rabbit (fragment)

C:Species: *Oryctolagus cuniculus* (domestic rabbit)

C>Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 29-Sep-1999

C:Accession: I46528

R:Payor, A.M.

J. Biol. Chem. 270, 5779-5785, 1995

A:Title: Sequence and functional characterization of a renal sodium/dicarboxylate cot

A:Reference number: I46528; MUID:95197598

A:Accession: 146528
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-593 <PAJ>
 A:Cross-references: EMBL:U12186; NID:q758383; PIDN:AAA9666.1; PID:q758384
 C:Superfamily: sodium/sulfate cotransporter

Query Match 82.5%; Score 33; DB 2; Length 593;
 Best Local Similarity 75.0%; Pred. No. 42;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GRVLVODG 8
 |||:|
 Db 367 GRVMSDG 374

RESULT 7

hypothetical protein T05E8.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T29621
 R:Mu, X.; Le, T.T.
 Submitted to the EMBL Data Library, April 1997

A:Description: The sequence of C. elegans cosmid T05E8.
 A:Reference number: Z20654
 A:Accession: T29621
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-893 <WUX>
 A:Cross-references: EMBL:U97014; PIDN:AAB52425.1; GSPDB:GNO0019; CESP:T05E8.1
 A:Experimental source: strain Bristol N2; clone T05E8
 C:Genetics:
 A:Gene: CESP:T05E8.1
 A:Map position: 1
 A:Introns: 22/3; 87/3; 185/3; 244/3; 297/3; 367/2; 404/3; 545/2; 673/3

Query Match 82.5%; Score 33; DB 2; Length 893;
 Best Local Similarity 62.5%; Pred. No. 65;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GRVLVODG 8
 |||:|
 Db 78 GRVLVODG 85

RESULT 8

2-demethylmenaquinone 2-C-methyltransferase (EC 2.1.1.-) HI0508 [similarity] - *Haemophilus*
 C:Species: *Haemophilus influenzae*
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Sep-2000
 C:Accession: G64153
 R:Flitschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
 ; Gocayne, J.D.; Scott, J.; Shirdley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fitchman, J.L.; Fuhman, J.L.; Geoghegan, N.S.M.
 Science 269, 496-512, 1995

A:Authors: Gocayne, J.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
 ; Fitchman, J.L.; Fuhman, J.L.; Geoghegan, N.S.M.
 A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
 A:Reference number: A64000; MUID:95350630
 A:Accession: G64153

A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA

A:Residues: 1-162 <TIGR>
 A:Cross-references: GB:U32732; GB:U42023; NID:q1573480; PIDN:AAC22166.1; PID:q1573489; T
 A:Note: best homolog was a hypothetical protein from *Escherichia coli*
 C:Keywords: methyltransferase

Query Match 80.0%; Score 32; DB 2; Length 162;
 Best Local Similarity 87.5%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GRVLVODG 8
 |||:|
 Db 58 GRVLVODG 65

RESULT 9

ferripyochelin-binding protein homolog b1400 - *Escherichia coli*
 C:Species: *Escherichia coli*
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 24-Sep-1999
 C:Accession: C64891
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; MUID:97426617
 A:Accession: C64891
 A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-196 <BLAT>
 A:Cross-references: GB:AE000237; GB:U00096; NID:q1787665; PIDN:AAC74482.1; PID:q17876
 A:Experimental source: strain K-12, substrain MG1655
 C:Superfamily: ferripyochelin binding protein

Query Match 80.0%; Score 32; DB 2; Length 196;
 Best Local Similarity 62.5%; Pred. No. 20;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GRVLVODG 8
 |||:|
 Db 48 GRVLVODG 55

RESULT 10

serine O-acetyltransferase (EC 2.3.1.30) cyse [validated] - *Bacillus subtilis*
 C:Species: *Bacillus subtilis*
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C:Accession: B53402; S66122; G69611
 R:Ogasawara, N.; Nakai, S.; Yoshikawa, H.
 DNA Res. 1, 1-14, 1994

A:Title: Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* ch
 A:Reference number: S65967; MUID:96051385
 A:Accession: S66122
 A:Status: nucleic acid sequence not shown; translation not shown

A:Residues: 1-217 <GAG>
 A:Molecule type: DNA
 A:Accession: B53402
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-217 <GAG>
 A:Cross-references: EMBL:D26185; NID:q467326; PIDN:BA05327.1; PID:q467481
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
 A:Authors: Foulger, D.; Foltz, C.; Fujita, M.; Fujita, Y.; Funai, S.; Gallizzi, A.; Gal
 leich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
 koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lartino
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
 Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Roche, B.; Roche, R.; Rose, M.; Sadle, Y.; Sato, T.; Scani
 A:Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiyuchi, J.; Sekowska, A.; Se
 akuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Togonni, A.; Tosato, V.; Uchlya

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033
A:Accession: G69611
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-217 <KUN>
A:Cross-references: GB:299104; GB:AL009126; NID:g2632267; PIDN:CAB11869.1; PID:g2632360
A:Experimental source: strain 168
C:Genetics:
A:Gene: cysE; cysA
A:Start codon: GTG
C:Function:
A:Description: EC 2.3.1.30 [validated; MUID:94171772]
A:Note: rate-limiting step in cysteine biosynthesis
C:Superfamily: *Bacillus* serine acetyltransferase; serine acetyltransferase homology
C:Keywords: acyltransferase; coenzyme A
F:9-169/Domain: serine acetyltransferase homology <SAR>

Query Match 80.0%; Score 32; DB 1; Length 217;
Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
|||:|:|
Db 166 GRVYVQNG 173

RESULT 11
C83267
Hypothetical protein PA3036 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83267
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
edman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,
., Lory, S.; Olson, M.V.
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
gen.
A:Reference number: A82950; MUID:20437337
A:Accession: C83267
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-310 <STO>
A:Cross-references: GB:AE004728; GB:AE004091; NID:g9949130; PIDN:AAG06424.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3036

Query Match 80.0%; Score 32; DB 2; Length 310;
Best Local Similarity 75.0%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
|||:|:|
Db 121 GRVLARDG 128

RESULT 12
JQ2398
Hypothetical 36.2K protein - citrus tatter leaf virus
N:Alternate names: ORF2 protein
C:Species: citrus tatter leaf virus
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C:Accession: JQ2398
R:Yoshikawa, N.; Imaizumi, M.; Takahashi, T.; Inouye, N.
J. Gen. Virol. 74, 2743-2747, 1993
A:Title: Striking similarities between the nucleotide sequence and genome organization o
f
A:Reference number: PQ0864; MUID:94103780
A:Accession: JQ2398

A:Molecule type: mRNA
A:Residues: 1-320 <YOS>
A:Cross-references: DDBJ:D16368; NID:g464150; PIDN:BAA03870.1; PID:g563956
A:Experimental source: strain Li-23

Query Match 80.0%; Score 32; DB 2; Length 320;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
|||:|:|
Db 107 GRVIVYDG 114

RESULT 13
B44059
36K protein - apple stem grooving virus (strain P-209)
C:Species: apple stem grooving virus, ASGV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 20-Jun-2000
C:Accession: B44059
R:Yoshikawa, N.; Sasaki, E.; Kato, M.; Takahashi, T.
Virology 191, 98-105, 1992
A:Title: The nucleotide sequence of apple stem grooving capillovirus genome.
A:Reference number: A44059; MUID:93033164
A:Accession: B44059
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-320 <YOS>
A:Cross-references: GB:D14995; NID:g303496; PIDN:BAA03640.1; PID:g285609

Query Match 80.0%; Score 32; DB 2; Length 320;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
|||:|:|
Db 107 GRVIVYDG 114

RESULT 14
B82197
Zona occludens toxin VC1458 [imported] - *Vibrio cholerae* (group O1 strain N16961)
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: B82197
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
Charlson, D.; Ermolaeva, M.D.; Yamthuevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20406833
A:Accession: B82197
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <HEI>
A:Cross-references: GB:AE004224; GB:AE003852; NID:g9655952; PIDN:AAF94615.1; GSPDB:GN
A:Experimental source: serogroup O1, strain N16961, biotype El Tor
C:Genetics:
A:Gene: VC1458
A:Map position: 1

Query Match 80.0%; Score 32; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
|||:|:|
Db 291 GRVLVODG 298

RESULT 15

A43864
Zonula occludens toxin - Vibrio cholerae
C:Species: Vibrio cholerae
C:Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C:Accession: A43864
R:Baudry, B.; Fasano, A.; Ketley, J.; Kaper, J.B.
Infect. Immun. 60, 428-434, 1992
A:Title: Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.
A:Reference number: A43864; MUID:92112300
A:Accession: A43864
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-399 <BAU>
A:Cross-references: CB:M83563; NID:g155314; PID:AAA27582.1; PID:g155315
A:Note: Sequence extracted from NCBI Backbone (NCBIN:77488, NCBIIP:77491)

Query Match 80.0%; Score 32; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 GRVLVQDG 8
||: |||
Db 291 GRLCVQDG 298

Search completed: June 13, 2001, 14:10:39
Job time: 140 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:21:40 ; Search time 74.44 Seconds
(without alignments)
3.681 Million cell updates/sec

Title: PCT-US01-05825A-4
Perfect score: 40
Sequence: 1 GRVLVQDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : SwlssProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	82.5	161	1 MENC_ECOLI	P32165 escherichia
2	33	82.5	161	1 STRD_STRGR	P08075 streptomyces
3	33	82.5	593	1 NDCL_RABIT	P028615 oryctolagus
4	32	80.0	162	1 MENC_HABIN	P44778 haemophilus
5	32	80.0	196	1 PAAY_ECOLI	P77181 escherichia
6	32	80.0	217	1 CYSE_BACSU	P06750 bacillus su
7	32	80.0	320	1 VPRF_ASCVP	P36698 apple stem
8	32	80.0	437	1 ZOT_VIRCH	P38442 vibrio chol
9	32	80.0	437	1 YF05_MERTH	P027549 methanobact
10	31	77.5	182	1 YCBO_ECOLI	P75855 escherichia
11	31	77.5	234	1 VG53_BPMUS	P05269 mycobacteri
12	31	77.5	235	1 VG53_BPMUS	P05269 mycobacteri
13	31	77.5	564	1 DPV4_RAT	P062951 rattus norv
14	31	77.5	572	1 DPV4_MOUSE	P035098 mus musculu
15	31	77.5	586	1 YAGM_RHTSN	P55469 rhizobium s
16	31	77.5	1317	1 RPOD_SYNY3	P73334 synechocyst
17	30	75.0	92	1 SR19_HALNI	P03921 halobacteri
18	30	75.0	101	1 DM71_ALIMI	P03960 alligator m
19	30	75.0	176	1 FRIM_SALSA	P49947 salmo salar
20	30	75.0	180	1 ULC1_HCVWA	P16741 human cytom
21	30	75.0	196	1 YEL9_YEAST	P40018 saccharomyc
22	30	75.0	205	1 DM71_TRASC	P57660 tracheomyx s
23	30	75.0	231	1 PYRE_PODAN	P08309 podospira a
24	30	75.0	306	1 SPEE_AERPE	P05602 aeropyrum p
25	30	75.0	308	1 SPD2_HYONI	P48659 hyoscyamus
26	30	75.0	311	1 DM71_CHICK	P09167 gallus gall
27	30	75.0	314	1 SPD2_NICSY	P48660 nicotiana s
28	30	75.0	315	1 SPD1_HYONI	P48658 hyoscyamus
29	30	75.0	316	1 SPD2_COPAR	P08214 coffea arab
30	30	75.0	317	1 SPD2_DATST	P05657 datura stra
31	30	75.0	323	1 SPD1_ORYSA	P09581 oryza sativ
32	30	75.0	334	1 SPD1_ARATH	P09203 arabidopsis
33	30	75.0	340	1 SPD2_ARATH	P48661 arabidopsis

34	30	75.0	342	1 SPD2_PEA	Q09210 pisum sativ
35	30	75.0	342	1 SPD2_LYCES	Q02645 lycopersico
36	30	75.0	434	1 PES_ERWCH	O51900 erwilia chr
37	30	75.0	434	1 GLYA_HYPME	P34895 hypomicrob
38	30	75.0	467	1 WCAM_SALTY	P26389 salmonella
39	30	75.0	610	1 YEM6_YEAST	P40024 saccharomyc
40	30	75.0	736	1 GEPH_RAT	P03555 rattus norv
41	30	75.0	831	1 DP01_THEFL	P30313 thermus aqu
42	30	75.0	832	1 DP01_THEAO	P19821 thermus aqu
43	30	75.0	834	1 DP01_THECA	P80194 thermus aqu
44	30	75.0	834	1 DP01_THETH	P52028 thermus aqu
45	30	75.0	1182	1 PTC2_MOUSE	O35595 mus musculu

ALIGNMENTS

RESULT 1	
MENC_ECOLI	STANDARD: PRT: 161 AA.
ID MENC_ECOLI	
AC P32165:	
DT 01-OCT-1993 (Rel. 27, Created)	
DT 01-OCT-1993 (Rel. 27, Last sequence update)	
DT 15-JUL-1998 (Rel. 36, Last annotation update)	
S-ADENOSYLMETHIONINE:2-DEMETHYLMENAOQUINONE METHYLTRANSFERASE	
DE (EC 2.1.1.-).	
GN MENC.	
OS Escherichia coli.	
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;	
OC Escherichia	
OX NCBI_TaxID=562;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=K12 / MG1655;	
RX MEDLINE=93347969; PubMed=8346018;	
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;	
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the	
RT region from 87.2 to 89.2 minutes."	
RL Nucleic Acids Res. 21:3391-3398(1993).	
CC -1- FUNCTION: CONVERTS DIMETHYLMENAOQUINONE (DMK) TO MENAOQUINONE (MK).	
CC -1- PATHWAY: LAST STEP IN MENAOQUINONE BIOSYNTHESIS.	
CC	
CC This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC the European Bioinformatics Institute. There are no restrictions on its	
CC use by non-profit institutions as long as its content is in no way	
CC modified and this statement is not removed. Usage by and for commercial	
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC or send an email to license@isb-sib.ch).	
CC	
DR EMBL: U56082; AAB01208.1; -	
DR EMBL: L19201; AAB03061.1; -	
DR EMBL: AE000467; AAC76911.1; -	
DR PIR: S40872; S40872.	
KR Ecogene: EGI1879; meng.	
DR Menaoquinone biosynthesis; Transferase; Methyltransferase.	
SO SEQUENCE 161 AA; 17360 MW; B30371B838DE21F8 CRC64;	
Query Match	82.5%; Score 33; DB 1; Length 161;
Best Local Similarity	82.5%; Pred. No. 3.5;
Matches	7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GRVLVQDG 8	
Db 58 GRVLVQDG 65	

```

RP SEQUENCE FROM N.A.
RC STRAIN-NEW ZEALAND WHITE; TISSUE-Kidney cortex;
RX MEDLINE=95197598; PubMed=7890707;
RA Pajor A.M.;
RT "Sequence and functional characterization of a renal
RL sodium/dicarboxylate cotransporter.";
RL J. Biol. Chem. 270:5779-5785(1995).
CC -1- FUNCTION: COTRANSPORT OF SODIUM AND DICARBOXYLATES SUCH AS
CC SUCCINATE AND CITRATE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- TISSUE SPECIFICITY: ABUNDANT IN KIDNEY AND SMALL INTESTINE.
CC -1- SIMILARITY: BELONGS TO THE NADc/P/PHO87 FAMILY OF TRANSPORTERS.
CC NADc SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: 012186; AAA99666.1; -
CC InterPro: IPR001898; -
CC Pfam: PRO0939; Na_Sulph_symp; 1.
CC PROSITE: PS01271; NA_SULFATE; 1.
DR Transport; Transmembrane; Sodium transport; Symport.
KW TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 121 141 POTENTIAL.
FT TRANSMEM 221 241 POTENTIAL.
FT TRANSMEM 270 290 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT TRANSMEM 369 389 POTENTIAL.
FT TRANSMEM 451 471 POTENTIAL.
FT TRANSMEM 485 505 POTENTIAL.
FT TRANSMEM 514 534 POTENTIAL.
FT TRANSMEM 543 563 POTENTIAL.
SQ SEQUENCE 593 AA; 65125 MW; 293BDB9J15D9E2BC CRC64;

Query Match 82.5%; Score 33; DB 1; Length 593;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVYVDG 8
|||:|
Db 367 GRVYVSDG 374

RESULT 4
MENG_HAEIN STANDARD; PRT; 162 AA.
AC P44738;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE S-ADENOSINETRIPHOSPHATE:2-DEOXYTHYMIDINADENINE METHYLTRANSFERASE
DE (EC 2.1.1.-.-).
GN MENG OR HI0508.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
OC Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shrivley R., Liu L.-I., Glodek A., Kelley J.M.,

```

```

RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT Influenzae Rd.":
RL Science 269:496-512(1995).
CC -1- FUNCTION: CONVERTS DIMETHYLMENAQUINONE (DMK) TO MENAQUINONE (MK).
CC -1- PATHWAY: LAST STEP IN MENAQUINONE BIOSYNTHESIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32732; AAC22166.1; -.
DR TIGR: H10508; -.
KW Menaquinone biosynthesis; Transferrase; Methyltransferase.
SQ SEQUENCE 162 AA; 17388 MW; B86E0E9BD2D8796A CRC64;

Query Match      80.0%; Score 32; DB 1; Length 162;
Best Local Similarity 87.5%; Pred. No. 5.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVDG 8
DB 58 GRVLVDG 65

RESULT 5
PAAY_ECOLI ID PAAY_ECOLI STANDARD: PRT; 196 AA.
AC P77181; O53020;
DT 01-OCT-2000 (rel. 40; Created)
DT 01-OCT-2000 (rel. 40; Last sequence update)
DT 01-OCT-2000 (rel. 40; Last annotation update)
DE PHENYLACETIC ACID DEGRADATION PROTEIN PAAY.
GN PAAY.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-9742617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-97251357; PubMed-9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
RA Oshima T., Saito N., Samped G., Seki Y., Sivasubraman S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RP SEQUENCE FROM N.A.
```

```

RC STRAIN-W / ATCC 11105;
RA Diaz E.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- PATHWAY: INVOLVED IN PHENYLACETIC ACID DEGRADATION.
CC -1- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000237; AAC74482.1; -.
DR EMBL: D90778; BAA15008.1; -.
DR EMBL: X97452; CAA66102.1; -.
DR Ecocore; E613747; paay.
DR InterPro; IPR001451; -.
DR Pfam; PF00132; hexapep; 1.
DR PROSITE; PS00101; HEXAPEP_TRANSFERASES; FALSE_NEG.
KW Transferase; Repeat.
FT VARIANT 75 75 G -> E (IN STRAIN W).
FT VARIANT 179 179 I -> V (IN STRAIN W).
FT VARIANT 182 182 G -> N (IN STRAIN W).
SQ SEQUENCE 196 AA; 21324 MW; FA3454F5AA0910DB CRC64;

Query Match      80.0%; Score 32; DB 1; Length 196;
Best Local Similarity 62.5%; Pred. No. 7.3;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLVDG 8
DB 48 GRVLVDG 55

RESULT 6
CYSE_BACSU ID CYSE_BACSU STANDARD: PRT; 217 AA.
AC 006750;
DT 01-JUN-1994 (rel. 29; Created)
DT 01-JUN-1994 (rel. 29; Last sequence update)
DT 01-FEB-1995 (rel. 31; Last annotation update)
DE SERINE ACETYLTRANSFERASE (EC 2.3.1.30) (SAT).
GN CYSE OR CYSA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9417172; PubMed-7510287;
RA Gagnon Y., Breton R., Putzer H., Pelchat M., Grunberg-Manago M.,
RA Lapointe J.;
RT "Clustering and co-transcription of the Bacillus subtilis genes
RT encoding the aminoacyl-tRNA synthetases specific for glutamate and
RT for cysteine and the first enzyme for cysteine biosynthesis.";
RL J. Biol. Chem. 269:7473-7482(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RX MEDLINE-96051385; PubMed-7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
RN [3]
RP SEQUENCE FROM N.A.
CC -1- CATALYTIC ACTIVITY: ACETYL-COA + L-SERINE -> COA + O-ACETYL-L-
CC SERINE.
CC -1- PATHWAY: CYSTEINE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
```

CC ACETYLTTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L14580; AAA21797.1; -
 DR EMBL: D26185; BAA05327.1; -
 DR EMBL: 299104; CAB1869.1; -
 DR PIR: B53402; B53402.
 DR HSSP: P01246; IBST.
 DR Subtilist: BG10155; cyse.
 DR InterPro: IPR001451; -
 DR Pfam: PF00132; hexapep; 2.
 DR PROSITE: PS00101; HEXAPEP; TRANSFERASES; 1.
 DR Transferrase: Acyltransferase; Cysteine biosynthesis; Repeat.
 SQ SEQUENCE 217 AA; 24143 MW; 1AEFA32FA1086D9D CRC64;

Query Match 80.0%; Score 32; DB 1; Length 217;
 Best Local Similarity 75.0%; Pred. No. 8.1;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GRVLVQDG 8
 Db 166 GRVYVQNG 173

RESULT 7
 ID VPRT_ASQVP STANDARD; PRT; 320 AA.
 AC P36696;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE PUTATIVE 36 KDA SERINE PROTEASE (EC 3.4.21.-) (ORF2).
 OS Apple stem grooving virus (Strain P-209) (ASGV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Capilliovirus.
 OX NCBI_TaxID=36402;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93033164; PubMed=1413530;
 RA Yoshikawa N., Sasaki E., Kato M., Takahashi T.;
 RT "The nucleotide sequence of Apple stem grooving capilliovirus genome.";
 RT Virology 191:98-105(1992).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S35.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D14995; BAA03640.1; -
 DR PIR: B44059; B44059.
 DR MEROPS: S35.001; -
 DR InterPro: IPR001815; -
 DR Pfam: PF02103; Peptidase_S35; 1.
 DR PRINTS: PRO00995; CAPILLOPTASE.
 DR Hydroxylase: Serine protease.
 FT ACT_SITE 144 144
 FT ACT_SITE 171 171
 FT ACT_SITE 199 199
 FT ACT_SITE 320 AA; 36136 MW; 656DFA9F505FB47A CRC64;
 SQ SEQUENCE

Query Match

80.0%; Score 32; DB 1; Length 320;

Best Local Similarity 75.0%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GRVLVQDG 8
 Db 107 GRVYVQNG 114

RESULT 8
 ID ZOT_VIBCH STANDARD; PRT; 399 AA.
 AC P38442; Q9L706; Q9R3V6;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE ZONA OCCUDENS TOXIN (ZONULAR OCCUDENS TOXIN).
 GN ZOT OR VC1458.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CLASSICAL INABA 569B;
 RX MEDLINE=92112300; PubMed=1730472;
 RA Baudry B., Fasano A., Ketter J., Kaper J.B.;
 RT "Cloning of a gene (zot) encoding a new toxin produced by Vibrio
 RT cholerae";
 RL Infect. Immun. 60:428-434(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KN1H002;
 RA Shin H.J., Park Y.C., Kim Y.C.;
 RT "Cloning and nucleotide sequence analysis of the virulence gene
 RT cassette from Vibrio cholerae KN1H002 isolated in Korea.";
 RL Misatunrhag Hoiji 35:205-210(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-0139-TOR OGAMA;
 RA Zhi-Yong H., Wei-Jie Z., Xiang-Fu W.;
 RT "Cloning and Expression of zot Gene from Vibrio cholerae";
 RL Submitted (Jan-1999) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EL TOR 86015 / SEROTYPE O1;
 RA Kan B., Liu Y.Q., Qi G.M., Gao S.Y.;
 RL Submitted (Jan-2000) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EL TOR N16961 / SEROTYPE O1;
 RA MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae";
 RL Nature 406:477-483(2000).
 RN [6]
 RP CHARACTERIZATION
 RA MEDLINE=91271365; PubMed=2052603;
 RA Fasano A., Baudry B., Pumpilin D.W., Wasserman S.S., Tall B.D.,
 RA Kelley J.M., Kaper J.B.;
 RT "Vibrio cholerae produces a second enterotoxin, which affects
 RT intestinal tight junctions";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5242-5246(1991).
 CC -1- FUNCTION: INCREASES THE PERMEABILITY OF THE SMALL INTESTINE MUCCOSA
 CC BY AFFECTING THE STRUCTURE OF INTERCELLULAR TIGHT JUNCTIONS
 CC (ZONULA OCCUDENS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M83563: AAA27582.1: -
DR EMBL: AF175708: AAD51558.1: -
DR EMBL: AF123049: AAD26854.1: -
DR EMBL: AF220606: AAF29547.1: -
DR EMBL: AE004224: AAF94615.1: -
DR PIR: A43864: A43864.
DR TIGR: VCL1458: -
KW Enterotoxin; Toxin.
FT VARIANT 45 45 M -> I (IN STRAIN 569B).
FT VARIANT 100 100 V -> A (IN STRAINS 569B AND 86015).
FT VARIANT 272 272 V -> A (IN STRAIN 569B).
FT VARIANT 281 281 V -> A (IN STRAIN 569B).
FT VARIANT 349 349 A -> S (IN STRAIN 86015).
FT VARIANT 381 381 K -> R (IN STRAIN 86015).
FT CONFLICT 386 399 IKTENDKKGLNSIF -> VKKEKEESIKSFL (IN REF.
FT 4).
SO SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;

Query Match 80.0%; Score 32; DB 1; Length 399;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
DB 291 GRVLVODG 298

RESULT 9
YF05_METTH STANDARD; PRT; 427 AA.
AC 027549;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL PROTEIN MTH1505.
GN MTH1505.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_Taxid=145262;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H;
RX MEDLINE=98037514; PubMed=9371463; Delonghery C., Lee H.-M., Dubois J.,
RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,
RA Aldridge T., Bashirzadeh R., Blakey D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pohler B., Olu D.,
RA Spadofora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell J., Prabhakar S.,
RA McDougall S., Shiner G., Goyal A., Pleitrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltam: functional analysis and comparative genomics."
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- SIMILARITY: BELONGS TO THE ATZ/TRZ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL: AE000910; AAB85980.1; -

DR InterPro: IPR002604: -
DR Pfam: PF01685: ATZ_TRZ; 1.
KW Hypothetical protein; Hydrolase.
SQ SEQUENCE 427 AA; 46441 MW; 439527DD142182CD CRC64;

Query Match 80.0%; Score 32; DB 1; Length 427;
Best Local Similarity 62.5%; Pred. No. 17;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
DB 397 GRVLVODG 404

RESULT 10
YCBQ_ECOLI STANDARD; PRT; 182 AA.
AC P75855;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL FIMBRIAL-LIKE PROTEIN IN PEPN-PYRD INTERGENIC REGION
DE PRECURSOR.
KW YCBQ.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1234-1238(1997).
[2]
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiiuchi T.;
RT "A 7.18-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:137-155(1996).
CC -1- SIMILARITY: BELONGS TO THE FIMA/PAPA FAMILY OF FIMBRIA PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL: AE000196: AAC74024.1; -
DR EMBL: D90732: BAA35693.1; -
DR EcoGene: EG13709: ycbQ.
DR InterPro: IPR000259: -
DR Pfam: PF00419: Fimbrin; 1.
KW Hypothetical protein; Fimbrin; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 182 HYPOTHETICAL FIMBRIAL-LIKE PROTEIN YCBQ.
SO SEQUENCE 182 AA; 19025 MW; A984F5CB2B68686C CRC64;

Query Match 77.5% Score 31; DB 1; Length 182;

Best Local Similarity 62.5%; Pred. No. 11;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLYVDG 8

Db 143 GNILLQDG 150

RESULT 11

VG53_BPML5 STANDARD; PRT; 234 AA.

AC Q05269;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DE 01-FEB-1994 (Rel. 28, Last annotation update)

GN GENE 53 PROTEIN (GP53).

OS Mycobacteriophage L5.

OC Viruses.

NCBI_TaxID=12376;

RP SEQUENCE FROM N.A.

RA MEDLINE=93211282; Pubmed=8459766;

RT "DNA sequence, structure and gene expression of mycobacteriophage L5;

RL Mol. Microbiol. 7:395-405(1993).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; 218946; CAA79429.1; -

DR PIR; S30998; S30998.

SQ SEQUENCE 234 AA; 26255 MW; 4205D93B006EE90D CRC64;

Query Match 77.5% Score 31; DB 1; Length 234;

Best Local Similarity 62.5%; Pred. No. 15;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLYVDG 8

Db 168 GRVLYVDG 175

RESULT 12

VG53_BPMD2 STANDARD; PRT; 235 AA.

AC O64243;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DE 15-DEC-1998 (Rel. 37, Last annotation update)

GN GENE 53 PROTEIN (GP53).

OS Mycobacteriophage D29.

OC Viruses.

NCBI_TaxID=28369;

RP SEQUENCE FROM N.A.

RA MEDLINE=96300335; Pubmed=9636706;

RT "Genome structure of mycobacteriophage D29: implications for phage evolution."

RL J. Mol. Biol. 279:143-164(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF022214; AAC18493.1; -

DR SEQUENCE 235 AA; 26339 MW; F7146CF6A12B935D CRC64;

RESULT 13

DPY4_RAT STANDARD; PRT; 564 AA.

AC Q62951;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DE 15-JUL-1998 (Rel. 36, Last annotation update)

GN DEHYDROXYRIMIDINASE RELATED PROTEIN-4 (DRP-4) (COLLAPSED RESPONSE MEDIATOR PROTEIN 3) (CRMP-3) (FRAGMENT).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;

RP SEQUENCE FROM N.A.

RA TISSUE=Brain;

RA MEDLINE=9644532; Pubmed=8815901;

RT "A family of rat CRMP genes is differentially expressed in the nervous system."

RL J. Neurosci. 16:6197-6207(1996).

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).

CC -1- DEVELOPMENTAL STAGE: EXPRESSED TRANSLATIONALLY IN DEVELOPING SPINAL CORD AND SELECTIVELY IN THE POSTNATAL CEREBELLUM.

CC -1- SIMILARITY: BELONGS TO THE DEHYDROXYRIMIDINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U52103; AAB03281.1; -

DR HSSP; P18316; IKRC.

DR InterPro: IPR002195; -

DR Pfam: PF00744; Dihydroorotase; 1.

FT NON_TER 1

SQ SEQUENCE 564 AA; 61085 MW; 42050891CC1A36D2 CRC64;

Query Match 77.5% Score 31; DB 1; Length 564;

Best Local Similarity 62.5%; Pred. No. 39;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLYVDG 8

Db 442 GRVLYVDG 449

RESULT 14

DPY4_MOUSE

DPY4_MOUSE

DPY4_MOUSE

DPY4_MOUSE

DPY4_MOUSE

DPY4_MOUSE

DPY4_MOUSE


```

ID DPV4_MOUSE STANDARD: PRT: 572 AA.
AC 035098: 008886:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DIHYDROPYRIMIDINASE RELATED PROTEIN-4 (DRP-4) (ULIP4 PROTEIN).
GN DPVSL4 OR ULIP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo.
RA Hamajima N., Kato Y., Kouvaki M., Wada Y., Sasaski M., Nonaka M.;
RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE=98314496; PubMed=9652388;
RA Byk T., Ozon S., Sobel A.;
RT "The Ulp family phosphoproteins -- common and specific properties.";
RL Eur. J. Biochem. 254:14-24(1998).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE DEHYDROPYRIMIDINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL: AB006715; BAA2188.1; -
DR EMBL: Y09079; CAA70299.1; -
DR MGD: MGI:1349764; Dpysl4.
DR InterPro: IPR002195; -
DR Pfam: PF00744; Dihydrocotarase; 1.
FT CONFLICT 125 126 ER -> DG (IN REF. 2).
FT CONFLICT 354 354 G -> V (IN REF. 2).
FT CONFLICT 420 420 F -> I (IN REF. 2).
SQ SEQUENCE 572 AA; 61961 MW; 37671129FC02C7AF CRC64;

Query Match 77.5%; Score 31; DB 1; Length 572;
Best Local Similarity 62.5%; Pred. No. 40;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLYODG 8
DB 450 GRVLYEDG 457

RESULT 15
V4GM_RHISN STANDARD: PRT: 586 AA.
AC P55469;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROBABLE ABC TRANSPORTER ATP-BINDING PROTEIN V4GM.
GN V4GM.
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Felley R., Bairoch A., Broughton W.J., Rosenthal A.,
Perret X.;

```

```

RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MSBA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL: AE000075; AAB91687.1; -
DR InterPro: IPR001140; -
DR InterPro: IPR001617; -
DR Pfam: PF00664; ABC_membrane; 1.
DR Pfam: PF00005; ABC_tran; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW Hypothetical protein; ATP-binding; Transport; Transmembrane; Plasmid.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 66 86 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT NP_BIND 379 386 ATP (POTENTIAL).
SQ SEQUENCE 586 AA; 64262 MW; 1095DFEB82620637 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 586;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLYODG 8
DB 565 GRVLSODG 572

```

Search completed: June 13, 2001, 14:21:41
Job time: 801 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:20:24 ; Search time 225.85 Seconds
(without alignments)
4.152 Million cell updates/sec

Title: PCT-US01-05825A-4
Perfect score: 40
Sequence: 1 GRVLVDG 8

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organellar:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-unclassified:*
13: SP-vertebrate:*
14: SP-virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	35	87.5	473	2	P96438	P96438 sinorhizobii
2	34	85.0	509	10	O9SAK4	O9SAK4 arabidopsis
3	34	85.0	528	10	O9SEK4	O9SEK4 arabidopsis
4	33	82.5	160	2	O9RW10	O9RW10 delinococcus
5	33	82.5	166	10	O9MR89	O9MR89 arabidopsis
6	33	82.5	245	2	O9L170	O9L170 streptomyces
7	33	82.5	482	1	O29120	O29120 archaeoglob
8	33	82.5	769	5	O9NMH8	O9NMH8 drosophila
9	33	82.5	893	5	O01596	O01596 caenorhabdi
10	32	80.0	320	14	P87586	P87586 citrus ta
11	32	80.0	320	14	O9YP11	O9YP11 citrus ta
12	32	80.0	323	2	O9L8F5	O9L8F5 vibrio mimi
13	32	80.0	399	2	O9R3V6	O9R3V6 vibrio chol
14	32	80.0	399	2	O9L706	O9L706 vibrio chol
15	32	80.0	590	10	O9SU18	O9SU18 arabidopsis
16	32	80.0	687	5	P91352	P91352 caenorhabdi
17	32	80.0	1767	5	O19346	O19346 caenorhabdi
18	31	77.5	117	2	O53026	O53026 nocardia co
19	31	77.5	123	2	O9RKN4	O9RKN4 streptomyces

20	31	77.5	168	4	O9NXM5	O9NXM5 homo sapien
21	31	77.5	273	2	O9WYP8	O9WYP8 thermocoga
22	31	77.5	309	1	O26846	O26846 methanobact
23	31	77.5	351	2	O44630	O44630 bruceella ab
24	31	77.5	447	10	O04511	O04511 arabidopsis
25	31	77.5	563	1	O28349	O28349 archaeoglob
26	31	77.5	575	10	O9M2G8	O9M2G8 arabidopsis
27	31	77.5	652	2	O9PEE1	O9PEE1 xylella fas
28	31	77.5	656	2	O07638	O07638 streptomyces
29	31	77.5	722	2	O9R226	O9R226 delinococcus
30	31	77.5	823	10	P93691	P93691 triticum ae
31	31	77.5	1100	5	O60948	O60948 trypanosoma
32	31	77.5	1392	10	O9SDC9	O9SDC9 oryza sativ
33	31	77.5	1600	10	O9XE34	O9XE34 oryza sativ
34	31	77.5	1665	5	O23464	O23464 caenorhabdi
35	30	75.0	101	13	O9PU60	O9PU60 alligator m
36	30	75.0	162	2	O9KPK1	O9KPK1 vibrio chol
37	30	75.0	163	2	O9SAU0	O9SAU0 pseudomonas
38	30	75.0	176	13	P79822	P79822 oncorhynch
39	30	75.0	176	13	P79823	P79823 oncorhynch
40	30	75.0	176	13	P79821	P79821 oncorhynch
41	30	75.0	193	2	O9KTN7	O9KTN7 vibrio chol
42	30	75.0	197	2	O9L0S1	O9L0S1 streptomyces
43	30	75.0	213	2	O9RI38	O9RI38 streptomyces
44	30	75.0	218	5	O09548	O09548 caenorhabdi
45	30	75.0	277	2	O9KX10	O9KX10 streptomyces

ALIGNMENTS

RESULT 1
ID P96438 PRELIMINARY; PRT; 473 AA.
AC P96438;
DT 01-MAY-1997 (TREMUREL. 03, Created)
DT 01-MAY-1997 (TREMUREL. 03, Last sequence update)
DT 01-MAY-2000 (TREMUREL. 13, Last annotation update)
DE EXPD2.
OS Sinorhizobium meliloti.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RM2011;
RX MEDLINE=97175570; PubMed=9023225;
RA Becker A., Rueberg S., Kuester H., Roxlau A.A., Keller M.,
RA Ivashina T., Cheng H., Walker G.C., Pehler A.,
RT "The 32-kilobase exp gene cluster of Rhizobium meliloti directing the
RT biosynthesis of galactoglucan: genetic organization and properties of
RT the encoded gene products.";
RL J. Bacteriol. 179:1375-1384(1997).
DR EMBL: Z79692; CAB01939.1; -;
DR INTERPRO: IPR002215; -;
DR PRAM: PR00529; HlyD; 1.
SQ SEQUENCE 473 AA: 51479 MW: C0C3B16BD9C23D85 CRC64;

Query Match 87.5% Score 35; DB 2; Length 473;
Best Local Similarity 75.0%; Pred. No. 40;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVLVDG 8
Db 107 GRVLVDG 114
RESULT 2
ID O9SAK4 PRELIMINARY; PRT; 509 AA.
AC O9SAK4;
DT 01-MAY-2000 (TREMUREL. 13, Created)

```

DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DE 01-OCT-2000 (TREMblrel. 15, last annotation update)
GN TRK14.14 PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Li J., Kremenetskaia I., Luros J., Lee J.M., Gonzalez A., Altati H.,
RA Araujo R., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,
RA Hultzer L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,
RA Davis R.W., Ecker J.R., Federpsiel N.A., Theologis A.;
RT Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Theologis A.;
RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Theologis A.;
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC007202; AAD30232.1; -.
DR HSSP: P20000; IAG8.
DR INTERPRO: IPR002086; -.
DR PFAM: PF00171; aldehyd. 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN.1.
DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN.1.
SQ SEQUENCE 509 AA; 54911 MW; 5259B64D5E316D2 CRC64;
QY 2 RVLVODG 8
Db 317 RVLVODG 323

Query Match 85.0%; Score 34; DB 10; Length 509;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVLVODG 8
Db 317 RVLVODG 323

RESULT 3
Q9SEK4 PRELIMINARY; PRT; 528 AA.
AC Q9SEK4;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
DE SUCCINIC SEMIALDEHYDE DEHYDROGENASE (EC 1.2.1.24).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Busch K., Fromm H.;
RT "Analysis of a purified recombinant plant mitochondrial succinic
RT semialdehyde dehydrogenase, an enzyme of the GABA shunt."
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: A111735; AAF23590.1; -.
DR HSSP: P20000; IAG8.
DR INTERPRO: IPR002086; -.
DR PFAM: PF00171; aldehyd. 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN.1.
DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN.1.

```

```

KW Oxidoreductase.
SQ SEQUENCE 528 AA; 56559 MW; A0DA95A3A592B5B9 CRC64;
QY 2 RVLVODG 8
Db 336 RVLVODG 342

Query Match 85.0%; Score 34; DB 10; Length 528;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVLVODG 8
Db 336 RVLVODG 342

RESULT 4
Q9RM10 PRELIMINARY; PRT; 160 AA.
AC Q9RM10;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DT 01-JUN-2000 (TREMblrel. 14, last annotation update)
DE S-ADENOSYLMETHIONINE:2-DEMETHYLMENAUINONE METHYLTRANSFERASE.
GN DR0859.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RA MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome Sequence of the Radioreistant Bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL: AE001940; AAF10437.1; -.
DR TIGR: DR0859; -.
KW transferase; Methyltransferase.
SQ SEQUENCE 160 AA; 16860 MW; 54D6F226CA19EC0E CRC64;
QY 1 GRVLVODG 8
Db 59 GRVLVODG 66

Query Match 82.5%; Score 33; DB 2; Length 160;
Best Local Similarity 87.5%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVLVODG 8
Db 59 GRVLVODG 66

RESULT 5
Q9M8R9 PRELIMINARY; PRT; 166 AA.
AC Q9M8R9;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
DE PUTATIVE S-ADENOSYLMETHIONINE:2-DEMETHYLMENAUINONE METHYLTRANSFERASE.
GN F13E7.29.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.;

```

RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC F13E7 genomic sequence."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC018363; AAF26983.1; -
 SQ SEQUENCE 166 AA; 17836 MW; D569E84FE28169F0 CRC64;

Query Match 82.5%; Score 33; DB 10; Length 166;
 Best Local Similarity 87.5%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
 |||||
 DB 64 GRVLVODG 71

RESULT 6
 O9L170 PRELIMINARY; PRT; 245 AA.
 AC O9L170;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE HYPOTHETICAL 26.2 KDA PROTEIN.
 GN SC3C11.02.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Redenbach M., Kleser H.M., Denaplatte D., Elchner A., Dullum J.,
 RA Klenash H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL M.J. Microbiol. 21:77-96(1996).
 DR EMBL: AL158060; CAB76308.1; -
 KM Hypothetical protein.
 SQ SEQUENCE 245 AA; 26169 MW; A327470E23F77612 CRC64;

Query Match 82.5%; Score 33; DB 2; Length 245;
 Best Local Similarity 87.5%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
 |||||
 DB 135 GRVLVODG 142

RESULT 7
 O29120 PRELIMINARY; PRT; 482 AA.
 AC O29120;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE 4-HYDROXYBUTYRATE COA TRANSFERASE (CAT2-1).
 GN AF1145.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.

OX NCBI_TaxID=2234;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kellavag A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
 RA Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus."
 RL Nature 390:364-370(1997).
 DR EMBL: AE001025; AAB90101.1; -
 DR TIGR: AF1145; -
 DR PRODOM: PD005956; -; 1.
 KW Hypothetical protein; Transferase.
 SQ SEQUENCE 482 AA; 54018 MW; 895ADBBD410E1CAF8 CRC64;

Query Match 82.5%; Score 33; DB 1; Length 482;
 Best Local Similarity 75.0%; Pred. No. 1;e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
 |||||
 DB 225 GRVLVODG 232

RESULT 8
 O9NHX8 PRELIMINARY; PRT; 769 AA.
 AC O9NHX8;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE C12.2 (FRAGMENT).
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Phylloidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Batterham P., Fair J.A., Bogwitz M., Grubor V., Pollock J.A.,
 RA Behan K.J., Nichols C.;
 RT "Molecular structure of the lozenge gene of Drosophila melanogaster."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF217651; AAF35310.1; -
 FT NON-TER 769
 SQ SEQUENCE 769 AA; 86699 MW; 3A4E3818BD251E29 CRC64;

Query Match 82.5%; Score 33; DB 5; Length 769;
 Best Local Similarity 87.5%; Pred. No. 1.8e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
 |||||
 DB 160 GRVLVODG 167

RESULT 9
 O01596 PRELIMINARY; PRT; 893 AA.
 AC O01596;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)

DT 01-JUL-1997 (TREMblrel. 04, last sequence update)
 DE 01-NOV-1998 (TREMblrel. 08, last annotation update)
 GN COSMID T05E8.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kerhaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierly-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans."
 RN Nature 368:32-38(1994).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Wu X., Le T.T.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U97014; AAB52425.1; -;
 SQ SEQUENCE 893 AA; 102488 MW; 91C7AE14694A0DA CRC64;
 QY 1 GRVLVODG 8
 DB 78 GRIVRDS 85
 Query Match 82.5%; Score 33; DB 5; Length 893;
 Best Local Similarity 62.5%; Pred. No. 2.1e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 RESULT 10
 ID P87586 PRELIMINARY; PRT; 320 AA.
 AC P87586;
 DT 01-MAY-1997 (TREMblrel. 03, Created)
 DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
 DE 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE CELL TO CELL MOVEMENT PROTEIN.
 OS Citrus tatter leaf virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Capilliovirus.
 OX NCBI_TaxID=33759;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94333644; PubMed=8073638;
 RA Ohlta K., Ito T., Kawai A., Namba S., Kusumi T., Tsuchizaki T.;
 RT "Nucleotide sequence of the 3'-terminal region of citrus tatter leaf
 RT virus RNA."
 RL Virus Genes 8:169-172(1994).
 DR EMBL; D14455; BAA03352.1; -;
 DR INTERPRO: IPR001815; -;
 DR PFAM: PF02103; Peptidase_S35; 1.
 DR PRINTS: PR00995; CAPILLOPTASE.
 SQ SEQUENCE 320 AA; 36144 MW; 097ABDD3C64FDA3 CRC64;

Query Match 80.0%; Score 32; DB 14; Length 320;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRVLVODG 8
 DB 107 GRVIVYDG 114
 Query Match 80.0%; Score 32; DB 14; Length 320;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 RESULT 12
 ID Q9L8F5 PRELIMINARY; PRT; 323 AA.
 AC Q9L8F5;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE ZOT (FRAGMENT).
 GN ZOT.
 OS Vibrio mimicus.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=674;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-pm5;
 RX MEDLINE=20143766; PubMed=10678967;
 RA Boyd E.F., Moyer K.E., Shi L., Waldor M.K.;
 RT "Infectious CTXphi and the Vibrio pathogenicity island prophage in
 RT vibrio mimicus: evidence for recent horizontal transfer between V.

RT mimicus and V. cholerae."
RL Infect. Immun. 68:1507-1513(2000).
DR EMBL: AF207857; AAF40142.1; -.
FT NON_TER 1
SQ SEQUENCE 323 AA; 36306 MW; 01C12DAE9B872C3B CRC64;

Query Match 80.0%; Score 32; DB 2; Length 323;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
DB 240 GRVLVODG 247

RESULT 13

O9R3V6 PRELIMINARY: PRT: 399 AA.
AC O9R3V6: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ZONULAR OCCLOUDENS TOXIN (ZONA OCCLOUDENS TOXIN).
CN ZOT OR VC1458.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KNH002:
RA Shin H.J., Park Y.C., Kim Y.C.;
RT "Cloning and nucleotide sequence analysis of the virulence gene
cassette from Vibrio cholerae KNIH002 isolated in Korea."
RL Miselimunang Ho1j1 35:205-210(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-O139-TOR OGANA:
RA Zhi-Yong H., Wei-Jie Z., Xiang-Fu W.;
RT "Cloning and Expression of zot Gene from Vibrio cholerae."
RL Submitted (Jan-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-EL TOR N16961 / SEROTYPE O1:
RA MEDLIN=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae."
RL Nature 406:477-483(2000).
DR EMBL: AF175708; AAD51358.1; -.
DR EMBL: AF123049; AAD26854.1; -.
DR EMBL: AE004224; AAF94615.1; -.
DR TIGR: VC1458; -.
SQ SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;

Query Match 80.0%; Score 32; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
DB 291 GRVLVODG 298

RESULT 14

O9L706
ID O9L706 PRELIMINARY: PRT: 399 AA.
AC O9L706: 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ZOT.
CN ZOT.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-86015;
RA Kan B., Liu Y.O., Qi G.M., Gao S.Y.;
RT "Vibrio cholerae nct-Ctxphi whole genome, include rstr(Rstr),
rsta(Rsta), rslb(Rslb), cep(Cep), orfU(Orfu), ace(Ace) and zot(zot)
genes."
RL Submitted (Jan-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF220606; AAF29547.1; -.
SQ SEQUENCE 399 AA; 44930 MW; CFE6A3DBC9E23BE1 CRC64;

Query Match 80.0%; Score 32; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
DB 291 GRVLVODG 298

RESULT 15

O9S018 PRELIMINARY: PRT: 590 AA.
AC O9S018: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 66.2 KDA PROTEIN.
GN T20K18.30 OR AT4G12680.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA Bancroft I., Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (Mar-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL049640; CAB40985.1; -.
DR EMBL: AL161534; CAB78310.1; -.
KW Hypothetical protein.
SQ SEQUENCE 590 AA; 66204 MW; 629A9681DC0CDAD3 CRC64;

Query Match 80.0%; Score 32; DB 10; Length 590;
Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
DB 393 GRILLRDG 400

Wed Jun 13 15:01:08 2001

pct-us01-05825a-4.rspt

Page 6

Search completed: June 13, 2001, 14:20:25
Job time: 725 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 13, 2001, 14:16:32 : Search time 118.55 Seconds
(without alignments)
1.296 Million cell updates/sec

Title: PCT-US01-05825A-4

Perfect score: 40

Sequence: 1 GRVLVODG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	80.0	400	2	US-08-624-601-8
2	31	77.5	117	1	US-08-499-215-3
3	30	75.0	79	4	US-09-025-151-7
4	30	75.0	310	2	US-08-484-956-88
5	30	75.0	310	2	US-08-757-653-88
6	30	75.0	315	2	US-08-484-956-91
7	30	75.0	315	2	US-08-757-653-91
8	30	75.0	320	2	US-08-757-653-163
9	30	75.0	320	2	US-08-823-516-61
10	30	75.0	320	3	US-08-759-038-102
11	30	75.0	320	3	US-08-758-314-102
12	30	75.0	322	2	US-08-484-956-89
13	30	75.0	322	2	US-08-757-653-89
14	30	75.0	528	2	US-08-484-956-90
15	30	75.0	528	2	US-08-757-653-90
16	30	75.0	548	2	US-08-484-956-86
17	30	75.0	548	2	US-08-757-653-86
18	30	75.0	695	2	US-08-484-956-87
19	30	75.0	695	2	US-08-757-653-87
20	30	75.0	830	1	US-07-977-434-6
21	30	75.0	830	1	US-08-458-819-6
22	30	75.0	830	5	PCT-US91-07035-6
23	30	75.0	831	1	US-08-073-384C-5
24	30	75.0	831	1	US-08-234-359A-5
25	30	75.0	831	1	US-08-483-043-5
26	30	75.0	831	1	US-08-481-238-5
27	30	75.0	831	2	US-08-471-066B-5

28	30	75.0	831	2	US-08-484-956-5	Sequence 5, Appl1
29	30	75.0	831	2	US-08-757-653-5	Sequence 5, Appl1
30	30	75.0	831	2	US-08-599-491-5	Sequence 5, Appl1
31	30	75.0	831	2	US-08-756-386-5	Sequence 5, Appl1
32	30	75.0	831	2	US-08-823-516-5	Sequence 5, Appl1
33	30	75.0	831	3	US-08-682-853A-5	Sequence 5, Appl1
34	30	75.0	831	3	US-08-759-038-5	Sequence 5, Appl1
35	30	75.0	831	3	US-08-758-314-5	Sequence 5, Appl1
36	30	75.0	832	1	US-07-977-434-2	Sequence 2, Appl1
37	30	75.0	832	1	US-08-156-020-2	Sequence 2, Appl1
38	30	75.0	832	1	US-08-156-020-4	Sequence 2, Appl1
39	30	75.0	832	1	US-08-156-020-6	Sequence 2, Appl1
40	30	75.0	832	1	US-08-156-020-8	Sequence 2, Appl1
41	30	75.0	832	1	US-08-156-020-10	Sequence 2, Appl1
42	30	75.0	832	1	US-08-073-384C-4	Sequence 4, Appl1
43	30	75.0	832	1	US-08-254-359A-4	Sequence 4, Appl1
44	30	75.0	832	1	US-08-483-043-4	Sequence 4, Appl1
45	30	75.0	832	1	US-08-458-819-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-624-601-8
Sequence 8, Application US/08624601
Patent No. 5882653
GENERAL INFORMATION:
APPLICANT: Kaper Dr., James B.
APPLICANT: Levine Dr., Myron M.
TITLE OF INVENTION: Vibrio cholerae O1 (CVD111) and non-O1
TITLE OF INVENTION: (CVD112 and CVD112R) serogroup vaccine strains, methods
of making same and products thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Spencer & Frank
STREET: 1100 New York Ave. N.W. Suite 300 East
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/624,601
FILING DATE: 08-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schneller Dr., John W.
REGISTRATION NUMBER: 26,031
REFERENCE/DOCKET NUMBER: BAC20019P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)414-4000
TELEFAX: (202)414-4040
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Vibrio cholerae
STRAIN: El Tor 7946
IMMEDIATE SOURCE:
CLONE: zot
US-08-624-601-8

Query Match 80.0%; Score 32; DB 2; Length 400;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GRVLVODG 8
Db 291 GRVLVODG 298

RESULT 2
US-08-499-215-3
; Sequence 3, Application US/08499215
; Patent No. 5612204
; GENERAL INFORMATION:
; APPLICANT: Saeki, Hisashi
; APPLICANT: Miura, Akira
; TITLE OF INVENTION: BIOLOGICAL DEGRADATIVE
; TITLE OF INVENTION: TREATMENT OF CHLORINE-SUBSTITUTED ETHYLENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue
; CITY: N.W.
; STATE: Washington D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/499,215
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; FILING DATE: 08-JUL-1994
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-499-215-3

Query Match 77.5%; Score 31; DB 1; Length 117;
Best Local Similarity 62.5%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 GRVLVODG 8
Db 95 GRVLVODG 102

RESULT 3
US-09-025-151-7
; Sequence 7, Application US/09025151
; Patent No. 6187535
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Fromont, Micheline
; APPLICANT: Rain, Jean-Christophe
; TITLE OF INVENTION: FAST AND EXHAUSTIVE METHOD FOR SELECTING A PREY
; TITLE OF INVENTION: POLYPEPTIDE INTERACTING WITH A BAIT POLYPEPTIDE OF
; TITLE OF INVENTION: INTERACTORS POLYPEPTIDES
; FILE REFERENCE: 03495-0164
; CURRENT APPLICATION NUMBER: US/09/025,151
; CURRENT FILING DATE: 1998-02-18

NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 7
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-025-151-7

Query Match 75.0%; Score 30; DB 4; Length 79;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 RVLVODG 8
Db 10 RVLVODG 16

RESULT 4
US-08-484-956-88
; Sequence 88, Application US/08484956
; Patent No. 5843634
; GENERAL INFORMATION:
; APPLICANT: DAHLBERG, JAMES E.
; APPLICANT: LYAMICHEV, VICTOR I.
; APPLICANT: BROW, MARY ANN D.
; APPLICANT: OLDENBURG, MARY C.
; APPLICANT: HEISTER, LAURA
; TITLE OF INVENTION: DETECTION OF p53 MUTATIONS
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,956
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/402,601
; FILING DATE: 09-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,164
; FILING DATE: 09-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/254,359
; FILING DATE: 06-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/073,384
; FILING DATE: 04-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/986,330
; FILING DATE: 07-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL J, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01801
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-956-88

Query Match 75.0%; Score 30; DB 2; Length 310;
Best Local Similarity 75.0%; Pred. No. 75;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
1111: 11
Db 13 GRVLVODG 20

RESULT 5
US-08-757-653-88
Sequence 88, Application US/08757653
Patent No. 5843669
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamlichev, Victor I.
APPLICANT: Lyamlichev, Natasha
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
TITLE OF INVENTION: Thermostable FEN-1 Endonucleases
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 88:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-757-653-88

Query Match 75.0%; Score 30; DB 2; Length 310;
Best Local Similarity 75.0%; Pred. No. 75;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
1111: 11
Db 13 GRVLVODG 20

RESULT 6
US-08-484-956-91
Sequence 91, Application US/08484956
Patent No. 5843654
GENERAL INFORMATION:
APPLICANT: DAHLBERG, JAMES E.

APPLICANT: LYAMICHEV, VICTOR I.
APPLICANT: BROW, MARY ANN D.
APPLICANT: OLDENBURG, MARY C.
APPLICANT: HEISLER, LAURA
TITLE OF INVENTION: DETECTION OF p53 MUTATIONS
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,956
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/402,601
FILING DATE: 09-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,164
FILING DATE: 09-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/254,359
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL J, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-956-91

Query Match 75.0%; Score 30; DB 2; Length 315;
Best Local Similarity 75.0%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVODG 8
1111: 11
Db 25 GRVLVODG 32

RESULT 7
US-08-757-653-91
Sequence 91, Application US/08757653
Patent No. 5843669
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamlichev, Victor I.
APPLICANT: Lyamlichev, Natasha
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
TITLE OF INVENTION: Thermostable FEN-1 Endonucleases

NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ. ID NO.: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-757-653-91

Query Match 75.0%; Score 30; DB 2; Length 315;
Best Local Similarity 75.0%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVLVDG 8
|||||
Db 25 GRVLVDG 32

RESULT 8
US-08-757-653-163
Sequence 163, Application US/08757653
Patent No. 5843669
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
TELECOMMUNICATION INFORMATION:

REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ. ID NO.: 163:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-757-653-163

Query Match 75.0%; Score 30; DB 2; Length 320;
Best Local Similarity 75.0%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVLVDG 8
|||||
Db 25 GRVLVDG 32

RESULT 9
US-08-823-516-61
Sequence 61, Application US/08823516
Patent No. 5994069
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Mast, Andrea L.
APPLICANT: Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-823-516-61

Query Match 75.0%; Score 30; DB 2; Length 320;
Best Local Similarity 75.0%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVDG 8
||||: ||
DB 25 GRVLVDG 32

RESULT 10
US-08-759-038-102
Sequence 102, Application US/08759038
Patent No. 6090543
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Brow, Mary Ann D.
APPLICANT: Dahlberg, James E.
TITLE OF INVENTION: Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,038
FILING DATE: 02-DEC-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02574
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-759-038-102

Query Match 75.0%; Score 30; DB 3; Length 320;
Best Local Similarity 75.0%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVDG 8
||||: ||
DB 25 GRVLVDG 32

RESULT 11
US-08-758-314-102
Sequence 102, Application US/08758314
Patent No. 6090606
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha
TITLE OF INVENTION: Improved Cleavage Agents
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,314
FILING DATE: 02-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02575
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-758-314-102

Query Match 75.0%; Score 30; DB 3; Length 320;
Best Local Similarity 75.0%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRVLVDG 8
||||: ||
DB 25 GRVLVDG 32

RESULT 12
US-08-484-956-89
Sequence 89, Application US/08484956
Patent No. 5843654

GENERAL INFORMATION:
APPLICANT: DAHLBERG, JAMES E.
APPLICANT: LYAMICHEV, VICTOR I.
APPLICANT: BROW, MARY ANN D.
APPLICANT: OLDBURG, MARY C.
APPLICANT: HEISLER, LAURA
TITLE OF INVENTION: DETECTION OF p53 MUTATIONS
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,956
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/402,601
FILING DATE: 09-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,164
FILING DATE: 09-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/254,359
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL J, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 322 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-956-89

Query Match 75.0%; Score 30; DB 2; Length 322;
Best Local Similarity 75.0%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVLVDG 8
DB 25 GRVLVDG 32

RESULT 13
US-08-757-653-89
Sequence 89, Application US/08757653
Patent No. 5843669
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha

TITLE OF INVENTION: Cleavage of Nucleic Acid Using
TITLE OF INVENTION: Thermostable FEN-1 Endonucleases
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 322 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-757-653-89

Query Match 75.0%; Score 30; DB 2; Length 322;
Best Local Similarity 75.0%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVLVDG 8
DB 25 GRVLVDG 32

RESULT 14
US-08-484-956-90
Sequence 90, Application US/08484956
Patent No. 5843654
GENERAL INFORMATION:
APPLICANT: DAHLBERG, JAMES E.
APPLICANT: LYAMICHEV, VICTOR I.
APPLICANT: BROW, MARY ANN D.
APPLICANT: OLDBURG, MARY C.
APPLICANT: HEISLER, LAURA
TITLE OF INVENTION: DETECTION OF p53 MUTATIONS
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,956
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/402,601
FILING DATE: 09-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,164
FILING DATE: 09-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/254,359
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL J, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 90:
SEQUENCE CHARACTERISTICS:
LENGTH: 528 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-956-90

Query Match 75.0%; Score 30; DB 2; Length 528;
Best Local Similarity 75.0%; Pred. NO. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Caps 0;

QY 1 GRVLVDG 8
DB 9 GRVLVDG 16

RESULT 15
US-08-757-653-90
Sequence 90 Application US/08757653
Patent No. 5843665
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha
TITLE OF INVENTION: Cleavage of Nucleic Acid Using
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 90:
SEQUENCE CHARACTERISTICS:
LENGTH: 528 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-757-653-90

Query Match 75.0%; Score 30; DB 2; Length 528;
Best Local Similarity 75.0%; Pred. NO. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Caps 0;

QY 1 GRVLVDG 8
DB 9 GRVLVDG 16

Search completed: June 13, 2001, 14:16:32
Job time: 493 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:14:32 : Search time 229.28 Seconds
(without alignments)
1.995 Million cell updates/sec

Title: PCT-US01-05825A-5

Perfect score: 46
Sequence: 1 GRLCVQPC 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_0401.*
2: /SIDS6/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SIDS6/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SIDS6/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SIDS6/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SIDS6/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SIDS6/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SIDS6/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SIDS6/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SIDS6/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SIDS6/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SIDS6/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SIDS6/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SIDS6/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SIDS6/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SIDS6/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SIDS6/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SIDS6/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SIDS6/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SIDS6/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SIDS6/gcgdata/geneseq/geneseq/AA1999.DAT.*
22: /SIDS6/gcgdata/geneseq/geneseq/AA2000.DAT.*
23: /SIDS6/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	8	21	Y79109
2	43	93.5	8	21	Y79105
3	39	84.8	8	21	Y79121
4	38	82.6	8	21	Y79110
5	38	82.6	8	21	Y79113
6	38	82.6	399	13	R20006
7	37	80.4	470	14	R34476
8	37	80.4	470	14	R34477
9	37	80.4	470	14	R34478
10	36	78.3	8	21	Y79111
11	36	78.3	8	21	Y79117

12	35	76.1	8	21	Y79106	Peptide antagonist
13	35	76.1	100	21	G17503	Arabidopsis thalia
14	35	76.1	129	21	G17502	Arabidopsis thalia
15	35	76.1	166	21	G17501	Arabidopsis thalia
16	35	76.1	236	21	B41465	Human OREF ORF1229
17	35	76.1	331	21	G41445	Arabidopsis thalia
18	35	76.1	878	20	Y00868	S. tuberosum isoam
19	35	76.1	942	20	Y00871	Original S. tubero
20	34	73.9	321	20	Y17300	Recombinant gp 90
21	34	73.9	901	21	B42494	Human OREF ORF2258
22	34	73.9	1291	16	R75201	Tyrosine phosphat
23	33	71.7	8	21	Y79107	Peptide antagonist
24	33	71.7	31	20	W88384	Human zneu1 partia
25	33	71.7	73	20	W88389	Human zneu1 partia
26	33	71.7	115	21	B41718	Human OREF ORF1482
27	33	71.7	153	21	B41638	Human OREF ORF1402
28	33	71.7	158	20	W88388	Human zneu1 partia
29	33	71.7	169	20	W88390	Human zneu1 partia
30	33	71.7	224	20	Y59870	Human normal uterus
31	33	71.7	247	21	Y52139	Human TANGO 125b (
32	33	71.7	254	20	W88382	Human neuro-growth
33	33	71.7	265	21	B42204	Human OREF ORF1968
34	33	71.7	273	20	Y41769	Human PRO213-1 pro
35	33	71.7	273	20	Y41770	Human PRO1330 prot
36	33	71.7	273	20	Y41771	Human PRO1449 prot
37	33	71.7	273	20	W88381	Human neuro-growth
38	33	71.7	273	21	B44325	Human PRO213-1 pro
39	33	71.7	273	21	B44326	Human PRO1449 prot
40	33	71.7	273	21	B44327	Human PRO1449 prot
41	33	71.7	273	21	B18673	Amino acid sequenc
42	33	71.7	273	21	B18674	Amino acid sequenc
43	33	71.7	273	21	B18675	Amino acid sequenc
44	33	71.7	273	21	B24042	Human PRO213 prote
45	33	71.7	273	21	B24043	Human PRO1330 prot

ALIGNMENTS

RESULT	1
ID	Y79109 standard; Peptide: 8 AA.
XX	Y79109;
AC	Y79109;
XX	05-JUN-2000 (first entry)
DT	
XX	
DE	Peptide antagonist of zonulin.
XX	
KW	Zonulin; antagonist; zonula occludens toxin receptor;
KW	blood-brain barrier; antiinflammatory; cerebroprotective;
KW	neuroprotective; dermatological; antitumor; antiviral;
KW	antibacterial; cytostatic; anti-HIV; vulnery; anti allergic;
KW	hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW	gastrointestinal inflammation; therapy.
XX	
OS	Synthetic.
XX	
PN	WO200007609-A1.
XX	
PD	17-FEB-2000.
XX	
PF	28-JUL-1999; 99WO-US16683.
XX	
PR	03-AUG-1998; 98US-0127815.
XX	
PA	(UYMA-) UNIV MARYLAND BALTIMORE.
XX	
PI	Fasano A;
XX	
DR	WPI; 2000-205565/18.
XX	
PT	New peptide antagonist of zonulin useful as antiinflammatory agent for

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 XX shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS Claim 1, Page 42; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulin, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

Sequence 8 AA:

Query Match 100.0%; Score 46; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRICVQPG 8
 DB 1 gricvpg 8

RESULT 2

ID Y79105 standard; Peptide; 8 AA.

XX Y79105;

DT 05-JUN-2000 (first entry)

XX Peptide antagonist of zonulin.

KW Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiulcer; antiviral;
 KW antibacterial; cytosstatic; anti-HIV; vulnerrary; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.

XX Synthetic.

OS WO200007609-A1.

PN 17-FEB-2000.

PD 28-JUL-1999; 99WO-US16683.

PR 03-AUG-1998; 98US-0127815.

PA (UYMA-) UNIV MARYLAND BALTIMORE.

PI Fasnano A;
 XX WPI; 2000-205565/18.

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 XX treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS Claim 1, Page 41; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulin, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

Sequence 8 AA:

Query Match 93.5%; Score 43; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRICVQPG 8
 DB 1 gricvpg 8

RESULT 3

ID Y79121 standard; Peptide; 8 AA.

XX Y79121;

DT 05-JUN-2000 (first entry)

XX Peptide antagonist of zonulin.

KW Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antiulcer; antiviral;
 KW antibacterial; cytosstatic; anti-HIV; vulnerrary; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.

XX Synthetic.

OS WO200007609-A1.

PN 17-FEB-2000.

PD 28-JUL-1999; 99WO-US16683.

XX 03-AUG-1998; 98US-0127815.
PR (UYMA-) UNIV MARYLAND BALTIMORE.
PA
XX
PI Fasano A;
XX WPI: 2000-205565/18.
DR
XX New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
XX
PS Claim 1; Page 46; 69pp; English.
XX
CC This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor. Yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.
CC
CC Sequence 8 AA:
SQ
Query Match 84.8%; Score 39; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. No. 3.2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GRICVQPG 8
DB 1 99lcwvqg 8
RESULT 4
Y79110
ID Y79110 standard; Peptide: 8 AA.
XX Y79110;
XX
XX 05-JUN-2000 (first entry)
XX
XX Peptide antagonist of zonulin.
XX
XX Zonulin; antagonist; zonula occludens toxin receptor;
KW blood-brain barrier; antiinflammatory; cerebroprotective;
KW neuroprotective; dermatological; antiulcer; antiviral;
KW antibacterial; cytostatic; anti-HIV; vulnerrary; antiallergic;
KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW gastrointestinal inflammation; therapy.
XX
XX Synthetic.
XX
XX OS

PN W0200007609-A1.
XX
XX 17-FEB-2000.
PD
XX 28-JUL-1999; 99WO-US16683.
XX
XX 03-AUG-1998; 98US-0127815.
XX
XX (UYMA-) UNIV MARYLAND BALTIMORE.
PA
XX Fasano A;
XX WPI: 2000-205565/18.
DR
XX New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
XX
PS Claim 1; Page 42; 69pp; English.
XX
CC This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor. Yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.
CC
CC Sequence 8 AA:
SQ
Query Match 82.6%; Score 38; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. No. 3.2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GRICVQPG 8
DB 1 99lcwvqg 8
RESULT 5
Y79113
ID Y79113 standard; Peptide: 8 AA.
XX Y79113;
XX
XX 05-JUN-2000 (first entry)
XX
XX Peptide antagonist of zonulin.
XX
XX Zonulin; antagonist; zonula occludens toxin receptor;
KW blood-brain barrier; antiinflammatory; cerebroprotective;
KW neuroprotective; dermatological; antiulcer; antiviral;
KW antibacterial; cytostatic; anti-HIV; vulnerrary; antiallergic;
KW

KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.
 OS Synthetic.
 XX WO200007609-A1.
 PN 17-FEB-2000.
 PD 28-JUL-1999; 99WO-US16683.
 PF 03-AUG-1998; 98US-0127815.
 PR (UYMA-) UNIV MARYLAND BALTIMORE.
 PA Fasano A;
 PI WPI: 2000-205565/18.
 DR New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS Claim 1; Page 43; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 SQ Sequence 8 AA:

Query Match 82.6%; Score 38; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRICVQPG 8
 II IIIII
 Db 1 gricvqpg 8

RESULT 6
 R20006
 ID R20006 standard; Protein: 399 AA.
 XX R20006;
 AC
 XX 31-MAR-1992 (first entry)
 DT
 XX zonula occludens toxin.
 DE

XX ZOT; cholera; vaccine; enterotoxin; diarrhoea.
 KW Vibrio cholerae.
 OS
 XX WO9118979-A.
 PN 12-DEC-1991.
 PD 05-JUN-1991; 91WO-US03812.
 PF 05-JUN-1990; 90US-0533315.
 PR (UYMA-) UNIV MARYLAND BALTI.
 PA Kaper JB, Baudry-Maurelli B, Fasano A;
 PI WPI: 1992-007465/01.
 DR N-PSDB; Q20185.
 XX New Vibrio cholerae strains - comprise restriction endonuclease
 PT fragment encoding toxin, used as vaccines against cholera
 PS Disclosure; Fig 18; 83pp; English.

CC The amino acid sequence is that of the zonula occludens toxin (ZOT).
 CC It may be responsible for diarrhoea in some strains of cholera and
 CC the ZOT gene or fragments of it are deleted from strains of Vibrio
 CC cholerae (V.c.) to be used as vaccines. These V.c. strains have 100%
 CC efficacy in protecting humans against subsequent infection with a
 CC strain of a similar serotype and avoid undesirable side effects such
 CC as diarrhoea, nausea and cramping. Cultures of these strains may be
 CC used for prodn. of vaccines against cholera.
 CC
 SQ Sequence 399 AA:

Query Match 82.6%; Score 38; DB 13; Length 399;
 Best Local Similarity 87.5%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRICVQPG 8
 II IIIII
 Db 291 gricvqpg 298

RESULT 7
 R34476
 ID R34476 standard; Protein: 470 AA.
 XX R34476;
 AC
 XX 30-JUL-1993 (first entry)
 DT
 XX Encoded by Hepatitis C virus clone JK4-A.
 DE HCV; non-A, non-B hepatitis virus; NANBV; liver disease;
 KW polymerase chain reaction; diagnostic method.
 XX Hepatitis C virus.
 OS
 XX JP05068562-A.
 PN 23-MAR-1993.
 PD 30-MAY-1991; 91JP-0153736.
 PF 30-MAY-1991; 91JP-0153736.
 PR (SANW) SANWA KAGAKU KENKYUSHO CO.
 PA WPI: 1993-130638/16.
 XX N-PSDB; Q40434.
 DR

xx DNA and cDNA of hepatitis C virus - useful as probes for
PT diagnosing HCV infection
xx
xx
PS Claim 4; Page 32-34; 44pp; Japanese.
xx
CC cDNA was prepared from HCV genomic RNA. Full-length clone Jk1-B
CC (9405 nucleotides long) and 14 shorter clones were isolated by PCR
CC amplification, including clone Jk4-A. Primer/probes derived from the
CC sequences of these clones can be used in diagnostic assays for HCV.
CC See Q40425-Q40439.
xx
SQ Sequence 470 AA;

Query Match 80.4%; Score 37; DB 14; Length 470;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRUCVOPG 8
DB 73 gracaqpg 80

RESULT 8
R34477
ID R34477 standard; Protein: 470 AA.
AC R34477;
XX
XX 30-JUL-1993 (first entry)
DE Encoded by Hepatitis C virus clone Jk4-B.
XX
XX HCV; non-A, non-B hepatitis virus; NANBH; liver disease;
KM polymerase chain reaction; diagnostic method.
XX
XX Hepatitis C virus.
XX
XX Jp05068562-A.
XX
XX 23-MAR-1993.
PD 30-MAY-1991; 91JP-0153736.
XX
XX 30-MAY-1991; 91JP-0153736.
PR
XX (SANM) SANMA KAGAKU KENKYUSHO CO.
PA
XX WPI: 1993-130638/16.
DR N-PSDB; Q40435.
XX
XX DNA and cDNA of hepatitis C virus - useful as probes for
PT diagnosing HCV infection
PT
XX
PS Claim 4; Page 34-36; 44pp; Japanese.
xx
CC cDNA was prepared from HCV genomic RNA. Full-length clone Jk1-B
CC (9405 nucleotides long) and 14 shorter clones were isolated by PCR
CC amplification, including clone Jk4-B. Primer/probes derived from the
CC sequences of these clones can be used in diagnostic assays for HCV.
CC See Q40425-Q40439.
xx
SQ Sequence 470 AA;

Query Match 80.4%; Score 37; DB 14; Length 470;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRUCVOPG 8
DB 73 gracaqpg 80

RESULT 9
R34478
ID R34478 standard; Protein: 470 AA.
XX
XX
AC R34478;
XX
XX 30-JUL-1993 (first entry)
DE Encoded by Hepatitis C virus clone Jk4-C.
XX
XX HCV; non-A, non-B hepatitis virus; NANBH; liver disease;
KM polymerase chain reaction; diagnostic method.
XX
XX Hepatitis C virus.
OS
XX
XX Jp05068562-A.
PN
XX 23-MAR-1993.
PD
XX
XX 30-MAY-1991; 91JP-0153736.
PF
XX 30-MAY-1991; 91JP-0153736.
PR
XX 30-MAY-1991; 91JP-0153736.
XX
XX (SANM) SANMA KAGAKU KENKYUSHO CO.
PA
XX WPI: 1993-130638/16.
DR N-PSDB; Q40436.
XX
XX DNA and cDNA of hepatitis C virus - useful as probes for
PT diagnosing HCV infection
PT
XX
PS Claim 4; Page 36-38; 44pp; Japanese.
XX
XX cDNA was prepared from HCV genomic RNA. Full-length clone Jk1-B
CC (9405 nucleotides long) and 14 shorter clones were isolated by PCR
CC amplification, including clone Jk4-C. Primer/probes derived from the
CC sequences of these clones can be used in diagnostic assays for HCV.
CC See Q40425-Q40439.
XX
XX
SQ Sequence 470 AA;

Query Match 80.4%; Score 37; DB 14; Length 470;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRUCVOPG 8
DB 73 gracaqpg 80

RESULT 10
Y79111
ID Y79111 standard; Peptide: 8 AA.
XX
XX Y79111;
AC
XX 05-JUN-2000 (first entry)
DR
XX
DE Peptide antagonist of zonulin.
XX
XX Zonulin; antagonist; zonula occludens toxin receptor;
KM blood-brain barrier; antiinflammatory; cerebroprotective;
KM neuroprotective; dermatological; antitumor; antiviral;
KM antibacterial; cytostatic; anti-HIV; vulnereary; antiallergic;
KM hypotensive; immunosuppressive; antiparasitic; vasotropic;
KM gastrointestinal inflammation; therapy.
XX
XX Synthetic.
OS
XX WO200007609-A1.

XX 17-FEB-2000.
 PD
 XX 28-JUL-1999; 99WO-US16683.
 PF
 XX 03-AUG-1998; 98US-0127815.
 PR
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA
 XX Fasano A;
 PI
 XX WPI; 2000-205565/18.
 DR
 XX
 PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS
 XX
 PS Claim 1; Page 43; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC celiac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 CC Sequence 8 AA:

QY 1 GRLGVOPG 8
 DB 1 grlivpgs 8
 Query Match 78.3%; Score 36; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 11
 ID Y79117 standard; Peptide; 8 AA.
 XX Y79117;
 AC
 XX 05-JUN-2000 (first entry)
 DT
 XX Peptide antagonist of zonulin.
 DE
 XX Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; anticancer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnerability; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotrophic;
 KM

KM gastrointestinal inflammation; therapy.
 XX
 XX Synthetic.
 OS
 XX WO200007609-A1.
 PN
 XX 17-FEB-2000.
 PD
 XX 28-JUL-1999; 99WO-US16683.
 PF
 XX 03-AUG-1998; 98US-0127815.
 PR
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA
 XX Fasano A;
 PI
 XX WPI; 2000-205565/18.
 DR
 XX
 PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS
 XX
 PS Claim 1; Page 45; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC celiac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 CC Sequence 8 AA:

QY 1 GRLGVOPG 8
 DB 1 ggvctvpgs 8
 Query Match 78.3%; Score 36; DB 21; Length 8;
 Best Local Similarity 75.0%; Pred. No. 3.2e+05;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RESULT 12
 ID Y79106 standard; Peptide; 8 AA.
 XX Y79106;
 AC
 XX 05-JUN-2000 (first entry)
 DT
 XX Peptide antagonist of zonulin.
 DE
 XX

XX Zonulin; antagonist; zonula occludens toxin receptor;
KW blood-brain barrier; antiinflammatory; cerebroprotective;
KW neuroprotective; dermatological; antitumor; antiviral;
KW antibacterial; cytostatic; anti-HIV; vulnerability; antiallergic;
KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW gastrointestinal inflammation; therapy.
XX Synthetic.
OS
XX
PN WO200007609-A1.
XX
PD 17-FEB-2000.
XX
XX 28-JUL-1999; 99WO-US16683.
PR 03-AUG-1998; 98US-0127815.
XX
XX (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PI Fasano A;
XX
XX WPI: 2000-205565/18.
XX
XX
PT New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
XX
PS Claim 1: Page 41; 69pp; English.
XX
XX This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical syntheses or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection.
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Kernerler's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.
XX
XX
SQ Sequence 8 AA:

Query Match 76.1%; Score 35; DB 21; Length 8;
Best Local Similarity 75.0%; Pred. No. 3.2e+05;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRUCVQPG 8
||:||||
Db 1 grvcvqdg 8

RESULT 13
G17503
ID G17503 standard; Protein: 100 AA.
XX
AC G17503:

XX 17-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 18546.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0122548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141847.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144874.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145182.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161355.
PR 26-OCT-1999; 99US-0161356.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.1%; Score 35; DB 21; Length 100;
Best Local Similarity 85.7%; Pred. NO. 21;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRLCVQP 7
DB 22 grlcvcvp 28

RESULT 14
ID G17502 standard; Protein; 129 AA.
G17502

AC G17502;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 18545.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;

Wed Jun 13 15:01:09 2001

```
KW      termination sequence.
XX      XX
OS      Arabidopsis thaliana.
XX      XX
PN      EPI033405-A2.
PD      06-SEP-2000.
XX      XX
PF      25-FEB-2000; 2000EP-0301439.
XX      XX
PR      25-FEB-1999; 99US-0121825.
PR      05-MAR-1999; 99US-0123180.
PR      09-MAR-1999; 99US-0123548.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      06-APR-1999; 99US-0128234.
PR      08-APR-1999; 99US-0128714.
PR      16-APR-1999; 99US-0129845.
PR      19-APR-1999; 99US-0130077.
PR      21-APR-1999; 99US-0130449.
PR      23-APR-1999; 99US-0130510.
PR      28-APR-1999; 99US-0130891.
PR      30-APR-1999; 99US-0131449.
PR      04-MAY-1999; 99US-0132407.
PR      05-MAY-1999; 99US-0132484.
PR      06-MAY-1999; 99US-0132485.
PR      06-MAY-1999; 99US-0132486.
PR      07-MAY-1999; 99US-0132487.
PR      11-MAY-1999; 99US-0132863.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134321.
PR      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.

PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      02-JUL-1999; 99US-0142154.
PR      06-JUL-1999; 99US-0142055.
PR      08-JUL-1999; 99US-0142390.
PR      09-JUL-1999; 99US-0142803.
PR      12-JUL-1999; 99US-0142820.
PR      13-JUL-1999; 99US-0142877.
PR      14-JUL-1999; 99US-0143542.
PR      15-JUL-1999; 99US-0143624.
PR      16-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      19-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144684.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      22-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145142.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      05-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
PR      09-AUG-1999; 99US-0147493.
PR      09-AUG-1999; 99US-0147935.
PR      10-AUG-1999; 99US-0148171.
PR      11-AUG-1999; 99US-0148319.
PR      12-AUG-1999; 99US-0148341.
PR      13-AUG-1999; 99US-0148565.
PR      13-AUG-1999; 99US-0148684.
PR      16-AUG-1999; 99US-0149368.
PR      17-AUG-1999; 99US-0149175.
PR      18-AUG-1999; 99US-0149426.
PR      20-AUG-1999; 99US-0149722.
PR      20-AUG-1999; 99US-0149723.
PR      20-AUG-1999; 99US-0149929.
PR      23-AUG-1999; 99US-0149902.
PR      23-AUG-1999; 99US-0149930.
PR      25-AUG-1999; 99US-0150566.
PR      26-AUG-1999; 99US-0150884.
PR      27-AUG-1999; 99US-0151065.
PR      27-AUG-1999; 99US-0151066.
PR      27-AUG-1999; 99US-0151080.
PR      30-AUG-1999; 99US-0151303.
PR      31-AUG-1999; 99US-0151348.
PR      01-SEP-1999; 99US-0151930.
PR      07-SEP-1999; 99US-0152263.
```

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155186.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159293.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match Best Local Similarity 76.1%; Score 35; DB 21; Length 129;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRLCVOP 7
Db 51 grlcvpd 57

RESULT 15

ID G17501 standard; Protein; 166 AA.

XX AC G17501;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 18544.

XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX EPI033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-1999; 99US-0121825.
XX 25-FEB-1999; 99US-0123180.
PR 05-MAR-1999; 99US-0123348.
PR 09-MAR-1999; 99US-0123578.
PR 23-MAR-1999; 99US-0126264.
PR 25-MAR-1999; 99US-0126785.
PR 29-MAR-1999; 99US-0127462.
PR 01-APR-1999; 99US-0128234.
PR 06-APR-1999; 99US-0128714.
PR 08-APR-1999; 99US-0129845.
PR 16-APR-1999; 99US-0130077.
PR 19-APR-1999; 99US-0130449.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135533.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140951.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.

Wed Jun 13 15:01:09 2001

```
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0152363.
PR 07-SEP-1999; 99US-0151930.
PR 31-AUG-1999; 99US-0151438.
PR 30-AUG-1999; 99US-0151303.
PR 13-SEP-1999; 99US-0153070.
PR 10-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.

PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159299.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.1%; Score 35; DB 21; Length 166;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GRLCVP 7
Db 88 grlcvp 94

Search completed: June 13, 2001, 14:14:32
Job time: 373 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:10:39 ; Search time 130.61 Seconds

(without alignments)
4.209 Million cell updates/sec

Title: PCT-US01-05825A-5

Perfect score: 46

Sequence: 1 GRICVOPG 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :
1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	87.0	447	2	T34992
2	38	82.6	399	2	B82197
3	38	82.6	399	2	A43864
4	37	80.4	505	2	H83196
5	37	80.4	782	2	S18032
6	35	76.1	508	1	B64204
7	35	76.1	593	2	S49525
8	34	73.9	96	2	D83228
9	34	73.9	255	2	G83014
10	34	73.9	859	1	VCLJEV
11	34	73.9	859	1	VCLJ22
12	34	73.9	859	1	VCLJEW
13	34	73.9	859	1	VCLJEI
14	34	73.9	859	1	VCLJEZ
15	34	73.9	859	1	VCLJE3
16	34	73.9	859	1	VCLJMS
17	34	73.9	1691	1	D54689
18	34	73.9	1894	2	C54689
19	33	71.7	245	2	S76632
20	33	71.7	379	2	A59180
21	33	71.7	391	2	C70972
22	33	71.7	417	2	T05207
23	33	69.6	508	2	S73430
24	32	69.6	508	2	T32867
25	32	69.6	226	2	T20645
26	32	69.6	302	2	T00480
27	32	69.6	345	2	D72536
28	32	69.6	358	2	T36415
29	32	69.6	482	2	C72254

30	32	69.6	494	2	F83199	probable carbhydr
31	32	69.6	498	2	JE0391	glycerol kinase (E
32	32	69.6	505	2	C82422	glycerol kinase VC
33	32	69.6	531	2	T50964	related to RC1 pr
34	32	69.6	539	2	I46470	estrogen dependent
35	32	69.6	614	2	B71551	probable s/t prote
36	32	69.6	614	2	H81703	conserved hypotet
37	32	69.6	619	2	D81556	conserved hypotet
38	32	69.6	619	2	A72114	s/t protein kinase
39	32	69.6	680	2	PN0510	integrin beta-3 ch
40	32	69.6	723	2	PN0509	integrin beta-3 ch
41	31.5	68.5	318	2	JC4963	metalloproteinase
42	31	67.4	92	2	D37057	epithelial cell gl
43	31	67.4	164	2	G64365	hypothetical prote
44	31	67.4	168	2	H69149	conserved hypotet
45	31	67.4	169	2	C75595	hypothetical prote

ALIGNMENTS

RESULT 1
T34992
probable lipoprotein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, August 1998
A:Reference number: Z21550
A:Accession: T34992
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-447 <OLIT>
A:Cross-references: EMBL:AL031182; PIDN:CA20169.1; GSPDB:GNO0070; SCOEDB:SC4A2.17C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC4A2.17C

Query Match 87.0%; Score 40; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRICVOP 7
DB 31 GRICVOP 37

RESULT 2
B82197
zona occludens toxin VC1458 [imported] - Vibrio cholerae (group O1 strain N16961)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: B82197
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833
A:Accession: B82197

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <HEI>
A:Cross-references: GB:AE004224; GB:AE003852; NID:9655952; PIDN:AAF94615.1; GSPDB:GN
C:Experimental source: serogroup O1, strain N16961; biotype El Tor
A:Genetics:
A:Gene: VC1458
A:Map position: 1

Query Match 82.6%; Score 38; DB 2; Length 399;

Best Local Similarity 87.5%; Pred. No. 8.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRICVOPG 8
DB 291 GRICVOPG 298

RESULT 3

A43864
zonula occludens toxin - Vibrio cholerae
C:Species: Vibrio cholerae
C>Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C:Accession: A43864
R:Baadry, B.; Fasanaro, A.; Kelley, J.; Kaper, J.B.
Infect. Immun. 60, 428-434, 1992
A:Title: Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.
A:Reference number: A43864; MUID:92112300
A:Accession: A43864
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-399 <PAU>
A:Cross-references: GB:M83563; NID:g155314; PIDN:AAA27582.1; PID:g155315
A:Note: sequence extracted from NCBI backbone (NCBIN:77488, NCBIP:77491)

Query Match 82.6%; Score 38; DB 2; Length 399;
Best Local Similarity 87.5%; Pred. No. 8.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRICVOPG 8
DB 291 GRICVOPG 298

RESULT 4

H83196
glycerol kinase PA3582 [Imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83196
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Brackman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Ladbis, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
A:Reference number: A82950; MUID:20437337
A:Accession: H83196
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-505 <STO>
A:Cross-references: GB:AE004779; GB:AE004091; NID:g9949735; PIDN:AG06970.1; GSPDB:GN001
A:Experimental source: strain PA01
A:Gene: gIPK; PA3582
C:Superfamily: xylokinnase

Query Match 80.4%; Score 37; DB 2; Length 505;
Best Local Similarity 62.5%; Pred. No. 17;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRICVOPG 8
DB 256 GOMCVERP 263

RESULT 5

S18032
genome polyprotein - hepatitis C virus (isolate Jk4) (fragment)
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
C:Species: hepatitis C virus
A:Variety: isolate Jk4

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C:Accession: S18032

R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.

submitted to the EMBL Data Library, September 1991

A:Description: Sequence analysis of putative structural regions of Hepatitis C virus

A:Reference number: S18029

A:Accession: S18032

A:Molecule type: genomic RNA

A:Residues: 1-782 <HON>

A:Cross-references: EMBL:X61594

A:Experimental source: isolate Jk4

C:Keywords: capsid protein; hepatitis C virus genome polyprotein

F:1191/Product: core protein; status predicted <MAT1>

F:192-383/Product: envelope protein 1 #status predicted <MAT2>

F:384-733/Product: NS1/E2 protein #status predicted <MAT3>

F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match 80.4%; Score 37; DB 2; Length 782;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRICVOPG 8
DB 73 GRACVOPG 80

RESULT 6

B64204
glycerol kinase (EC 2.7.1.30) - Mycoplasma genitalium
C:Species: Mycoplasma genitalium
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 07-Dec-1999
C:Accession: B64204
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.M.; Funtmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, C.A.; Venter, J.C.
Science 270, 397-403, 1995
A:Title: The minimal gene complement of Mycoplasma genitalium.
A:Reference number: A64200; MUID:96026346
A:Accession: B64204
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-508 <TIGR>
A:Cross-references: GB:U39682; GB:I43967; NID:g1045702; PID:g1045709; TIGR:MG038
A:Experimental source: strain G-37
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: xylokinnase
C:Keywords: phosphotransferase

Query Match 76.1%; Score 35; DB 1; Length 508;
Best Local Similarity 62.5%; Pred. No. 41;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRICVOPG 8
DB 258 GOLCTEPG 265

RESULT 7

S49525
glycoprotein G - simian herpesvirus B
N:Alternate names: US4 protein homolog
C:Species: simian herpesvirus B
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999
R:Slomka, M.J.; Brown, D.W.
submitted to the EMBL Data Library, October 1994
A:Description: Complete nucleotide sequence of simian herpes B virus glycoprotein G
A:Reference number: S49525
A:Accession: S49525

A:Molecule type: DNA
A:Residues: 1-593 <STO>
A:Cross-references: EMBL:Z46266; NID:9560495; PIDN:CAA86431.1; PID:9560496
A:Experimental source: Isolate Cyno 2
C:Keywords: glycoprotein

Query Match 76.1%; Score 35; DB 2; Length 593;
Best Local Similarity 75.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRLOCOPG 8
|||
Db 248 GRRCVSPG 255

RESULT 8

D83228 hypothetical protein PA3338 [Imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: D83228
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337
A:Accession: D83228

A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-96 <STO>
A:Cross-references: GB:AE004756; GB:AE004091; NID:99949466; PIDN:AMG06726.1; GSPDB:GN001

A:Experimental source: strain PA01
C:Genetics:

A:Gene: PA3338

Query Match 73.9%; Score 34; DB 2; Length 96;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRLOCOPG 8
|||
Db 38 GRVCLEPG 45

RESULT 9

G83014 Probable nuclease PA5048 [Imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: G83014
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337
A:Accession: G83014

A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-235 <STO>
A:Cross-references: GB:AE004918; GB:AE004091; NID:99951336; PIDN:AMG08433.1; GSPDB:GN001

A:Experimental source: strain PA01
C:Genetics:

A:Gene: PA5048

Query Match 73.9%; Score 34; DB 2; Length 255;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRLOCOPG 8
|||
Db 84 GRLOCOPG 91

RESULT 10

env polyprotein precursor - equine infectious anemia virus

N:Contains: coat protein gp45; coat protein gp90
C:Species: equine infectious anemia virus

A>Note: host Equus caballus (domestic horse)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 12-Apr-1996

C:Accession: A25610; B25610
R:Rushlow, K.; Olsen, K.; Stiegler, G.; Payne, S.L.; Montelaro, R.C.; Issel, C.J.

Virology 155, 309-321, 1986
A:Title: Lentivirus genomic organization: the complete nucleotide sequence of the env

A:Reference number: A25610; MUID:87071653
A:Accession: A25610

A:Molecule type: DNA
A:Residues: 1-859 <RUS>

C:Genetics:

A:Gene: env
C:Superfamily: equine infectious anemia virus env polyprotein

C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-444/Product: coat protein gp90 #status predicted <GPP>
F:445-859/Product: coat protein gp45 #status predicted <GGP>

F:446-472/Domain: transmembrane #status predicted <TMN>
F:617-636/Domain: transmembrane #status predicted <TMN>

F:40,112,141,148,186,214,233,244,340,368,399,406,411,483,490,550,557,752/Binding site

Query Match 73.9%; Score 34; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVOP 7
|||||
Db 265 RLCVOP 270

RESULT 11

VCLJ22 env polyprotein precursor - equine infectious anemia virus (strain CL22)

N:Alternate names: coat polyprotein
N:Contains: coat protein gp45; coat protein gp90

C:Species: equine infectious anemia virus
A>Note: host Equus caballus (domestic horse)

C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 16-Jul-1999
C:Accession: C41991

R:Perry, S.T.; Flaherty, M.T.; Kelley, M.J.; Clabough, D.L.; Tronick, S.R.; Coggins,

J. Virology 66, 4085-4097, 1992
A:Title: The surface envelope protein gene region of equine infectious anemia virus 1

A:Reference number: A41991; MUID:92292230
A:Accession: C41991

A:Molecule type: DNA
A:Residues: 1-859 <PBR>

A:Cross-references: GB:M87581; NID:9290627; PIDN:AAA43005.1; PID:9290630
C:Genetics:

A:Gene: env
C:Superfamily: equine infectious anemia virus env polyprotein

C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-859/Product: env polyprotein #status predicted <ENV>
F:23-444/Product: coat protein gp90 #status predicted <GPP>

F:445-859/Product: coat protein gp45 #status predicted <GGP>
F:446-472/Region: hydrophobic

F:615-636/Domain: transmembrane #status predicted <TMN>
F:40,112,141,148,186,214,233,244,340,368,399,406,411,483,490,550,557/Binding site: ca

Query Match 73.9%; Score 34; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVOP 7
|||||

Db 265 RLCVOP 270

RESULT 12

VCLJEM

env polyprotein precursor (clone 1369) - equine infectious anemia virus
N:Contains: coat protein gp45; coat protein gp90
C:Species: equine infectious anemia virus
A:Note: host Equus caballus (domestic horse)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 07-Nov-1997
C:Accession: C27842
R:Kawakami, T.; Sherman, L.; Dahlberg, J.; Gazit, A.; Yaniv, A.; Tronick, S.R.; Aaronson
Virology 158, 300-312, 1987
A:Title: Nucleotide sequence analysis of equine infectious anemia virus proviral DNA.
A:Reference number: A27842; MUID:87236196
A:Accession: C27842
A:Molecule type: DNA
A:Residues: 1-859 <RAW>
A:Cross-references: GB:M16575; NID:9323836
C:Genetics:
A:Gene: env
C:Superfamily: equine infectious anemia virus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-444/Product: coat protein gp90 #status predicted <CP1>
F:445-859/Product: coat protein gp45 #status predicted <CP2>
F:446-472/Domain: transmembrane #status predicted <TM1>
F:517-636/Domain: transmembrane #status predicted <TM2>
F:40,112,141,148,186,214,233,244,340,368,399,406,411,483,490,550,557,752/Binding site: C

Query Match 73.9%; Score 34; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVOP 7
|||||

Db 265 RLCVOP 270

RESULT 13

VCLJEL

env polyprotein precursor (clone P3.2-1) - equine infectious anemia virus
N:Contains: coat protein gp45; coat protein gp90
C:Species: equine infectious anemia virus
A:Note: host Equus caballus (domestic horse)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C:Accession: A34027
R:Payne, S.L.; Fang, F.D.; Liu, C.P.; Dhruva, B.R.; Rwambo, P.; Issel, C.J.; Montelaro,
Virology 161, 321-331, 1987
A:Title: Antigenic variation and lentivirus persistence: variations in envelope gene seq
A:Reference number: A34027; MUID:88072070
A:Accession: A34027
A:Molecule type: genomic RNA
A:Residues: 1-859 <PAV>
A:Cross-references: GB:M18385; NID:9323830; PIDN:AAA66407.1; PID:9323831
C:Genetics:
A:Gene: env
C:Superfamily: equine infectious anemia virus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-444/Product: coat protein gp90 #status predicted <CP1>
F:75-93/Domain: transmembrane #status predicted <TM1>
F:445-859/Product: coat protein gp45 #status predicted <CP2>
F:446-462/Domain: transmembrane #status predicted <TM2>
F:614-636/Domain: transmembrane #status predicted <TM3>
F:787-807/Domain: transmembrane #status predicted <TM4>
F:816-835/Domain: transmembrane #status predicted <TM5>
F:40,112,141,148,186,214,233,244,340,368,399,406,411,422,483,490,550,557,752/Binding site: C

Query Match 73.9%; Score 34; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVOP 7
|||||

Db 265 RLCVOP 270

RESULT 14

VCLJEE

env polyprotein precursor (clone P3.2-2) - equine infectious anemia virus
N:Contains: coat protein gp45; coat protein gp90
C:Species: equine infectious anemia virus
A:Note: host Equus caballus (domestic horse)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C:Accession: B34027
R:Payne, S.L.; Fang, F.D.; Liu, C.P.; Dhruva, B.R.; Rwambo, P.; Issel, C.J.; Montelaro,
Virology 161, 321-331, 1987
A:Title: Antigenic variation and lentivirus persistence: variations in envelope gene
A:Reference number: A34027; MUID:88072070
A:Accession: B34027
A:Molecule type: genomic RNA
A:Residues: 1-859 <PAV>
A:Cross-references: GB:M18386; NID:9323832; PIDN:AAA66408.1; PID:9323833
C:Genetics:
A:Gene: env
C:Superfamily: equine infectious anemia virus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-444/Product: coat protein gp90 #status predicted <CP1>
F:75-93/Domain: transmembrane #status predicted <TM1>
F:445-859/Product: coat protein gp45 #status predicted <CP2>
F:446-462/Domain: transmembrane #status predicted <TM2>
F:614-636/Domain: transmembrane #status predicted <TM3>
F:787-807/Domain: transmembrane #status predicted <TM4>
F:816-835/Domain: transmembrane #status predicted <TM5>
F:40,112,141,148,184,201,214,233,244,282,313,340,346,368,399,406,411,422,483,490,550,

Query Match 73.9%; Score 34; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVOP 7
|||||

Db 265 RLCVOP 270

RESULT 15

VCLJEE

env polyprotein precursor (clone P3.2-3) - equine infectious anemia virus
N:Contains: coat protein gp45; coat protein gp90
C:Species: equine infectious anemia virus
A:Note: host Equus caballus (domestic horse)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C:Accession: C34027
R:Payne, S.L.; Fang, F.D.; Liu, C.P.; Dhruva, B.R.; Rwambo, P.; Issel, C.J.; Montelaro,
Virology 161, 321-331, 1987
A:Title: Antigenic variation and lentivirus persistence: variations in envelope gene
A:Reference number: A34027; MUID:88072070
A:Accession: C34027
A:Molecule type: genomic RNA
A:Residues: 1-859 <PAV>
A:Cross-references: GB:M18387; NID:9323834; PIDN:AAA66409.1; PID:9323835
C:Genetics:
A:Gene: env
C:Superfamily: equine infectious anemia virus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-444/Product: coat protein gp90 #status predicted <CP1>

F:75-93/Domain: transmembrane #status predicted <TM1>
F:445-859/Product: coat protein gp45 #status predicted <CP2>
F:446-462/Domain: transmembrane #status predicted <TM2>
F:614-636/Domain: transmembrane #status predicted <TM3>
F:787-807/Domain: transmembrane #status predicted <TM4>
F:816-835/Domain: transmembrane #status predicted <TM5>
F:40,112,141,148,186,214,233,244,340,368,399,406,411,422,483,490,550,557,752/Binding site

Query Match 73.9%; Score 34; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;
Qy 2 RLCVQP 7
|||
Db 265 RLCVQP 270

Search completed: June 13, 2001, 14:10:39
Job time: 140 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:21:41 ; Search time 74.44 Seconds
(without alignments)
3.681 Million cell updates/sec

Title: PCT-US01-05825A-5
Perfect score: 46
Sequence: 1 GRICVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	82.6	399	1 ZOT_VIBCH	P38442 vibrio chol
2	37	80.4	505	1 GLPK_PSEAE	O61390 pseudomonas
3	35	76.1	331	1 GL2M_ARATH	O24430 arabidopsis
4	34	76.1	508	1 GLPK_MYCGE	P47284 mycoplasma
5	34	73.9	859	1 ENV_EIAY1	P22427 equine infe
6	34	73.9	859	1 ENV_EIAY2	P22428 equine infe
7	34	73.9	859	1 ENV_EIAY3	P22429 equine infe
8	34	73.9	859	1 ENV_EIAY4	P11306 equine infe
9	34	73.9	859	1 ENV_EIAY5	P12541 equine infe
10	34	73.9	859	1 ENV_EIAY6	P16081 equine infe
11	34	73.9	859	1 ENV_EIAY7	P06752 equine infe
12	33	71.7	327	1 O6A1_HUMAN	O95222 homo sapien
13	33	71.7	379	1 WIF1_HUMAN	O95222 homo sapien
14	33	71.7	379	1 WIF1_MOUSE	O95222 homo sapien
15	33	71.7	508	1 GLPK_MYCPN	P5064 mycoplasma
16	32	69.6	142	1 VLE4_RHIME	O97174 rhizobium m
17	32	69.6	496	1 GLK2_THEMA	O97164 thermotoga
18	32	69.6	339	1 OGP_SHEEP	O28542 ovis aries
19	32	69.6	636	1 P73_HUMAN	O15350 homo sapien
20	32	69.6	637	1 P73_CERAE	O95222 ceratophyc
21	32	69.6	787	1 ITB3_MOUSE	O54890 mus musculu
22	32	69.6	164	1 Y527_MERJA	O57729 methanococ
23	31	67.4	179	1 Y281_MERJA	O57729 methanococ
24	31	67.4	272	1 PSB0_STINE	P55221 synethococ
25	31	67.4	325	1 YP60_MYCTU	O50728 mycobacteri
26	31	67.4	355	1 SUR6_MOUSE	P70279 mus musculu
27	31	67.4	447	1 GNT1_RABIT	P71115 cytcoclagus
28	31	67.4	471	1 IAAG_MAIZE	O41819 zea mays (m
29	31	67.4	501	1 PHR1_SINAI	P40115 sinapis alb
30	31	67.4	547	1 Y073_CABEL	O09316 caenorhabdi
31	31	67.4	570	1 DCPY_NEUCR	P43287 neurospora
32	31	67.4	720	1 DCOR_HAEIN	P44317 haemophilus
33	31	67.4	731	1 DCOR_ECOLI	P11169 escherichia

34	31	67.4	732	1 DCOS_ECOLI	P24169 escherichia
35	31	67.4	760	1 SM4A_MOUSE	O62178 mus musculu
36	31	67.4	788	1 ITB3_HUMAN	P05106 homo sapien
37	31	67.4	900	1 AXN1_HUMAN	O15169 homo sapien
38	31	67.4	907	1 VGLB_HCMWT	P13201 human cytom
39	31	67.4	968	1 PKD2_HUMAN	O13563 homo sapien
40	31	67.4	992	1 AXN1_MOUSE	O35625 mus musculu
41	31	67.4	1021	1 Y1B8_CABEL	P46582 caenorhabdi
42	31	67.4	2907	1 FBN2_MOUSE	O61555 mus musculu
43	30	65.2	257	1 IOD1_RAT	P24389 ratius norv
44	30	65.2	278	1 EUTJ_ECOLI	P77277 escherichia
45	30	65.2	279	1 EUTJ_SALTY	P41794 salmonella

ALIGNMENTS

RESULT	ID	SEQUENCE	STANDARD	PRT	399 AA.
1	ZOT_VIBCH	P38442: Q9L706; Q9R3V6;			
AC	01-OCT-1994 (Rel. 30, Created)				
DT	01-OCT-2000 (Rel. 40, Last sequence update)				
DE	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	ZONA OCCIDENTIS TOXIN (ZONULAR OCCIDENTIS TOXIN).				
GN	ZOT OR VC1458.				
OS	Vibrio cholerae.				
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.				
OX	NCBI_Taxid=666;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CLASSICAL INABA 5699;				
KA	MEDLINE=92113300; PubMed=1730472;				
RA	Baudry B., Fasano A., Kelley J., Kaper J.B.;				
RT	"Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.";				
RL	Infect. Immun. 60:428-434(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=KN1H002;				
RA	Shin H.J., Park Y.C., Kim Y.C.;				
RT	"Cloning and nucleotide sequence analysis of the virulence gene cassette from Vibrio cholerae KN1H002 isolated in Korea.";				
RL	Mislimurthig Hoji 35:205-210(1999).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=O139-TOR OGAWA;				
RA	Zhi-Yong H., Wei-Jie Z., Xiang-Fu W.;				
RT	"Cloning and Expression of zot Gene from Vibrio cholerae.";				
RL	Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=EL TOR 86015 / SEROTYPE O1;				
RA	Kan B., Liu Y.O., Qi G.M., Gao S.Y.;				
RT	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.				
RL	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=EL TOR N16961 / SEROTYPE O1;				
RA	Medline=20406833; PubMed=10952301;				
RT	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,				
RA	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,				
RT	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,				
RA	Ernst A.M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,				
RT	McDonald L., Ueberback T., Fleischmann R.D., Nierman W.C., White O.,				
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,				
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.";				
RL	Nature 406:477-483(2000).				
RN	[6]				
RP	CHARACTERIZATION.				
RA	Medline=91271365; PubMed=2052603;				
RT	Fasano A., Baudry B., Pumpin D.W., Wasserman S.S., Tall B.D.,				

RA Kestley J.M., Kaper J.B.;
 RT "Vibrio cholerae produces a second enterotoxin, which affects
 RL intestinal tight junctions."
 CC Proc. Natl. Acad. Sci. U.S.A. 88:5242-5246(1991)
 CC -1- FUNCTION: INCREASES THE PERMEABILITY OF THE SMALL INTESTINE MUCOSA
 CC BY AFFECTING THE STRUCTURE OF INTERCELLULAR TIGHT JUNCTIONS
 CC (ZONULA OCCUDENS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: M83563; AAA27582.1; -
 CC DR EMBL: AF175708; AAD51358.1; -
 CC DR EMBL: AF123049; AAD56854.1; -
 CC DR EMBL: AF220506; AAF29547.1; -
 CC DR EMBL: AE004224; AAF94615.1; -
 CC DR PIR: A43864; A43864.
 CC DR TIGR: VC1458; -
 CC Enterotoxin; Toxin.
 CC KW VARIANT 45 45 M -> I (IN STRAIN 569B).
 CC FT VARIANT 100 100 V -> A (IN STRAINS 569B AND 86015).
 CC FT VARIANT 272 272 V -> A (IN STRAIN 569B).
 CC FT VARIANT 281 281 V -> A (IN STRAIN 569B).
 CC FT VARIANT 349 349 A -> S (IN STRAIN 86015).
 CC FT VARIANT 381 381 K -> R (IN STRAIN 86015).
 CC FT CONFLICT 386 399 IKTENDKKGLNSIF -> VKKEEESIIKNSFL (IN REF.
 CC FT SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;
 CC
 CC Query Match 82.6%; Score 38; DB 1; Length 399;
 CC Best Local Similarity 87.5%; Pred. No. 3.4;
 CC Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 GRLCYOPG 8
 CC DB 291 GRLCYOPG 298
 CC
 CC RESULT 2
 CC GLPK_PSEAE STANDARD: PRT: 505 AA.
 CC AC 051350;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-OCT-2000 (Rel. 40, Last sequence update)
 CC DT 01-OCT-2000 (Rel. 40, Last annotation update)
 CC DE GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE)
 CC GN GLPK OR PA3582.
 CC OS Pseudomonas aeruginosa.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC OC Pseudomonas.
 CC OX NCBI_TaxID=287;
 CC OX [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=PAOI;
 CC RC STRAIN-ATCC 15692 / PAOI;
 CC RX MEDLINE=97286544; PubMed=9141691;
 CC RA Schweizer H.P., Jump R., Po C.;
 CC RT "Structure and gene-polyptide relationships of the region encoding
 CC RT glycerol diffusion facilitator (glpf) and glycerol kinase (glpk) of
 CC RT Pseudomonas aeruginosa."
 CC RT Microbiology 143:1287-1297(1997).
 CC RL [2]
 CC RN
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=PAOI;
 CC RC MEDLINE=20437337; PubMed=10984043;
 CC RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 CC RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kass A., Lattig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Salzer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
 RT opportunistic pathogen".
 CC Nature 406:959-964(2000).
 CC
 CC -1- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND
 CC METABOLISM.
 CC -1- CATALYTIC ACTIVITY: ATP + GLYCEROL = ADP + GLYCEROL 3-PHOSPHATE.
 CC -1- PATHWAY: RATE-LIMITING STEP IN GLYCEROL UTILIZATION.
 CC -1- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCONOKINASE /
 CC GLYCEROKINASE / XYLOKINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: U49666; AAB57804.1; -
 CC DR EMBL: AE004779; AAG06970.1; -
 CC DR HSP: P08859; IGLB.
 CC DR InterPro: IPR000577; -
 CC DR Pfam: PF00370; PGGY. 1.
 CC DR PROSITE: PS00445; PGGY_KINASES. 2; 1.
 CC DR PROSITE: PS00933; PGGY_KINASES. 1; 1.
 CC KW Glycerol metabolism; Transferase; Kinase; ATP-binding.
 CC FT NP_BIND 158 170 ATP (PROBABLE).
 CC FT CONFLICT 75 75 H -> R (IN REF. 1).
 CC FT CONFLICT 104 104 A -> V (IN REF. 1).
 CC FT CONFLICT 109 109 C -> R (IN REF. 1).
 CC FT SEQUENCE 505 AA; 55966 MW; CC63A9AF8ABC752 CRC64;
 CC
 CC Query Match 80.4%; Score 37; DB 1; Length 505;
 CC Best Local Similarity 62.5%; Pred. No. 6.5;
 CC Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 GRLCYOPG 8
 CC DB 256 GOMCVERP 263
 CC
 CC RESULT 3
 CC GLZM_ARATH STANDARD: PRT: 331 AA.
 CC ID GLZM_ARATH
 CC AC 024495; 024494;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
 CC DE HYDROXYACYLGLUTATHIONE HYDROLASE, MITOCHONDRIAL PRECURSOR (EC 3.1.2.6)
 CC GN (GLYOXALASE II) (GLX II).
 CC GN GLX2-1.
 CC OS Arabidopsis thaliana (Mouse-ear cress).
 CC OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 CC OC Brassicales; Brassicaceae; Arabidopsids.
 CC OX NCBI_TaxID=3702;
 CC OX [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=CV. MASSILEWSKIIJA;
 CC RC MEDLINE=98009983; PubMed=9349270;
 CC RA Malt M.K., Krishnasamy S., Owen H.A., Makaroff C.A.;
 CC RT "Molecular characterization of glyoxalase II from Arabidopsis
 CC RT thaliana".
 CC RT Plant Mol. Biol. 35:471-481(1997).
 CC RL
 CC -1- FUNCTION: THIOLESTERASE THAT CATALYZES THE HYDROLYSIS OF S-D-
 CC LACTOYL-GLUTATHIONE TO FORM GLUTATHIONE AND D-LACTIC ACID.
 CC -1- CATALYTIC ACTIVITY: (S)-(-2-HYDROXYACYL)GLUTATHIONE + H(2)O =
 CC GLUTATHIONE + A 2-HYDROXY ACID ANION.

```

CC -1- COFACTOR: BINDS TWO ZINC IONS (BY SIMILARITY).
CC -1- PATHWAY: GLYOXAL PATHWAY.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -1- SIMILARITY: BELONGS TO THE GLYOXALASE II FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U09028; AAC49865.1; -.
CC DR EMBL: U09027; AAC49865.1; -.
CC DR InterPro: IPR001279; -.
CC KM Hydrolase; Zinc; Mitochondrion; Transalt peptide.
CC FT TRANSIT 1 76 MITOCHONDRION (POTENTIAL).
CC FT METAL 131 131 ZINC 1 (BY SIMILARITY).
CC FT METAL 133 133 ZINC 1 (BY SIMILARITY).
CC FT METAL 135 135 ZINC 2 (BY SIMILARITY).
CC FT METAL 136 136 ZINC 2 (BY SIMILARITY).
CC FT METAL 189 189 ZINC 1 (BY SIMILARITY).
CC FT METAL 208 208 ZINC 1 AND 2 (BY SIMILARITY).
CC FT METAL 159 159 H -> D (IN AAC49865).
CC FT CONFLICT 159 159
CC SQ SEQUENCE 331 AA; 36522 MW; 640E6892014DC24 CRC64;

Query Match 76.1%; Score 35; DB 1; Length 331;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRUCVOPG 8
   1:11111
Db 24 GOLCVMPG 31

RESULT 4
GLPK_MYCGE STANDARD; PRT; 508 AA.
AC P47284;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE)
DE (GLYCEROKINASE) (CK).
GN GLPK OR MG038.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fitchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
RA Nguyen D.T., Uutterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lueder T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A., III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL Science 270:397-403(1995).
CC -1- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND
CC METABOLISM.
CC -1- CATALYTIC ACTIVITY: ATP + GLYCEROL = ADP + GLYCEROL 3-PHOSPHATE.
CC -1- PATHWAY: RATE-LIMITING STEP IN GLYCEROL UTILIZATION.
CC -1- SIMILARITY: BELONGS TO THE FUCCOKINASE / GLUCONOKINASE /
CC GLYCEROKINASE / XYLUOKINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U39683; AAC71254.1; -.
CC DR HSSP; P08859; IGLB.
CC DR TIGR; MG038; -.
CC DR InterPro: IPR000577; -.
CC DR Pfam; PF00370; FGGY; 1.
CC DR PROSITE; PS00445; FGGY_KINASES_2; 1.
CC DR PROSITE; PS00933; FGGY_KINASES_1; 1.
CC KM GlyceroL metabolism; Transferrase; Kinase; ATP-binding.
CC FT NP_BIND 157 169 ATP (PROBABLE).
CC SQ SEQUENCE 508 AA; 56901 MW; 82033D7076D27CB5 CRC64;

Query Match 76.1%; Score 35; DB 1; Length 508;
Best Local Similarity 62.5%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRUCVOPG 8
   1:11111
Db 258 GOLCTEPG 265

RESULT 5
ENV_EIAY1 STANDARD; PRT; 859 AA.
AC P2427;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: COAT PROTEIN
DE GP90; COAT PROTEIN GP45].
GN ENV.
OS Equine infectious anemia virus (clone p3.2-1) (EIAV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11666;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88072070; PubMed=2825406;
RA Payne S.L., Fang F.D., Liu C.P., Dhruva B.R., Rambo P., Issel C.J.,
RA Montelaro R.C.;
RT "Antigenic variation and lentivirus persistence: variations in
RT envelope gene sequences during EIAV infection resemble changes
RT reported for sequential isolates of HIV."
RL Virology 161:321-331(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M18385; AAA66407.1; -.
CC DR PIR; A34027; VCLJEL.
CC DR InterPro: IPR001027; -.
CC DR InterPro: IPR001361; -.
CC DR Pfam; PF01045; EIAV_GP45; 1.
CC DR Pfam; PF00971; EIAV_GP90; 1.
CC KM Coat protein; Glycoprotein; Polypeptide; Transmembrane; Signal.
CC FT SIGNAL 1 22
CC FT CHAIN 23 859 ENV POLYPROTEIN.
CC FT CHAIN 23 444 COAT PROTEIN GP90.
CC FT CHAIN 445 859 COAT PROTEIN GP45.
CC FT TRANSMEM 75 93
CC FT TRANSMEM 446 462 POTENTIAL.
CC FT TRANSMEM 614 636 POTENTIAL.

```

```
FT TRANSMEM 787 807 POTENTIAL.
FT TRANSMEM 816 835 POTENTIAL.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 490 490 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 557 557 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 859 AA; 97140 MW; F4A0C071396DA867 CRC64;
```

```
Query Match
Best Local Similarity 73.9%; Score 34; DB 1; Length 859;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RLCVOP 7
Db 265 RLCVOP 270
```

```
RESULT 6
ID ENV_ETIAV2 STANDARD; PRT; 859 AA.
AC P22428;
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: COAT PROTEIN
DE GP90; COAT PROTEIN GP45].
GN ENV.
OS Equine infectious anemia virus (clone P3.2-2) (EIAV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID-11667;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88072070; PubMed-2825406;
RA Payne S.L., Fang F.D., Liu C.P., Dhruva B.R., Rambo P., Issel C.J.,
RA Montelaro R.C.;
RT "Antigenic variation and lentivirus persistence: variations in
RT envelope gene sequences during EIAV infection resemble changes
RT reported for sequential isolates of HIV.";
RL Virology 161:321-331(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M18386; AAA66408.1; -
CC PIR: B34027; VCLD2.
CC InterPro: IPR001027; -
CC DR InterPro: IPR001361; -
CC Pfam: PF01045; EIAV.GP45; 1.
CC Pfam: PF00971; EIAV.GP90; 1.
CC Coat protein; Glycoprotein; Polyprotein; Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 859 ENV POLYPROTEIN.
```

```
FT CHAIN 23 444 COAT PROTEIN GP90.
FT CHAIN 445 859 COAT PROTEIN GP45.
FT TRANSMEM 75 93 POTENTIAL.
FT TRANSMEM 446 462 POTENTIAL.
FT TRANSMEM 614 636 POTENTIAL.
FT TRANSMEM 787 807 POTENTIAL.
FT TRANSMEM 816 835 POTENTIAL.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 346 346 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 490 490 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 557 557 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 859 AA; 97188 MW; D86E4E1712E39B32 CRC64;
```

```
Query Match
Best Local Similarity 73.9%; Score 34; DB 1; Length 859;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RLCVOP 7
Db 265 RLCVOP 270
```

```
RESULT 7
ID ENV_ETIAV3 STANDARD; PRT; 859 AA.
AC P22429;
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: COAT PROTEIN
DE GP90; COAT PROTEIN GP45].
GN ENV.
OS Equine infectious anemia virus (clone P3.2-3) (EIAV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID-11668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88072070; PubMed-2825406;
RA Payne S.L., Fang F.D., Liu C.P., Dhruva B.R., Rambo P., Issel C.J.,
RA Montelaro R.C.;
RT "Antigenic variation and lentivirus persistence: variations in
RT envelope gene sequences during EIAV infection resemble changes
RT reported for sequential isolates of HIV.";
RL Virology 161:321-331(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
```

```
DR EMBL: M18387; AAA66409.1; -
DR PIR: C34027; VCLJE3.
DR InterPro: IPR001027; -
DR InterPro: IPR001361; -
DR Pfam: PF01045; EIAV_GP45; 1.
DR Pfam: PF00971; EIAV_GP90; 1.
KW Coat protein; Glycoprotein; Polypeptide; Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 859
FT CHAIN 23 444
FT CHAIN 445 859
FT TRANSMEM 75 93
FT TRANSMEM 446 462
FT TRANSMEM 614 636
FT TRANSMEM 787 807
FT TRANSMEM 816 835
FT CARBOHYD 40 40
FT CARBOHYD 112 112
FT CARBOHYD 141 141
FT CARBOHYD 148 148
FT CARBOHYD 186 186
FT CARBOHYD 214 214
FT CARBOHYD 233 233
FT CARBOHYD 244 244
FT CARBOHYD 340 340
FT CARBOHYD 368 368
FT CARBOHYD 399 399
FT CARBOHYD 406 406
FT CARBOHYD 411 411
FT CARBOHYD 422 422
FT CARBOHYD 430 483
FT CARBOHYD 483 490
FT CARBOHYD 550 550
FT CARBOHYD 557 557
FT CARBOHYD 752 752
SQ SEQUENCE 859 AA; 97066 MW; 982A9F5A1D8F4AD CRC64;

Query Match 73.9%; Score 34; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVOP 7
Db 265 RLCVOP 270

RESULT 8
ENV_EIAV9 STANDARD: PRT; 859 AA.
AC P11306;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: COAT PROTEIN
DE GP90; COAT PROTEIN GP45].
GN ENV.
OS Equine infectious anemia virus (clone 1369) (EIAV).
OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.
OX NCBI_TaxID=11670;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87236196; PubMed=3035786;
RA Kawakami T., Sherman L., Dahlberg J., Gazit A., Yaniv A.,
RA Tronick S.R., Aaronson S.A.;
RT "Nucleotide sequence analysis of equine infectious anemia virus
RT proviral DNA."
RL Virology 158:300-312(1987).
RN [2]
RP REVISIONS TO N-TERMINUS.
RA Tronick S.R.;
RL Submitted (NOV-1987) to the EMBL/GenBank/DBJ databases.
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M16575; AAB59863.1; -
DR PIR: C27842; VCLJEW.
DR InterPro: IPR001027; -
DR InterPro: IPR001361; -
DR Pfam: PF01045; EIAV_GP45; 1.
DR Pfam: PF00971; EIAV_GP90; 1.
KW Coat protein; Glycoprotein; Polypeptide; Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 859
FT CHAIN 23 444
FT CHAIN 445 859
FT TRANSMEM 75 93
FT TRANSMEM 446 462
FT TRANSMEM 614 636
FT TRANSMEM 787 807
FT TRANSMEM 816 835
FT CARBOHYD 40 40
FT CARBOHYD 112 112
FT CARBOHYD 141 141
FT CARBOHYD 148 148
FT CARBOHYD 186 186
FT CARBOHYD 214 214
FT CARBOHYD 233 233
FT CARBOHYD 244 244
FT CARBOHYD 340 340
FT CARBOHYD 368 368
FT CARBOHYD 399 399
FT CARBOHYD 406 406
FT CARBOHYD 411 411
FT CARBOHYD 483 483
FT CARBOHYD 490 490
FT CARBOHYD 550 550
FT CARBOHYD 557 557
FT CARBOHYD 752 752
SQ SEQUENCE 859 AA; 97113 MW; 4BAED8518CD4F364 CRC64;

Query Match 73.9%; Score 34; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVOP 7
Db 265 RLCVOP 270

RESULT 9
ENV_EIAVC STANDARD: PRT; 859 AA.
AC P32541;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: COAT PROTEIN
DE GP90; COAT PROTEIN GP45].
GN ENV.
OS Equine infectious anemia virus (clone C122) (EIAV).
OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.
OX NCBI_TaxID=31675;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92292230; PubMed=1318398;
RA Perry S.T., Flaherty M.T., Kelley M.J., Clabough D.L., Tronick S.R.,
RA Coggins L., Whetter L., Lengel C.R., Fuller F.;
RT "The surface envelope protein gene region of equine infectious anemia
```


[1]
RN SEQUENCE FROM N.A.
RX MEDLINE-87071653; PubMed-2431539;
RA Rueslow K., Olsen K., Stieglar G., Payne S.L., Montelaro R.C.,
Issel C.J.;
RT "Lentivirus genomic organization: the complete nucleotide sequence of
the env gene region of equine infectious anemia virus.";
RL Virology 155:309-321(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M16575; AAB59863.1; ALT_SEQ.
DR PIR: A25610; VCLJEV.
DR InterPro: IPR001027; -.
DR InterPro: IPR001361; -.
DR Pfam: PF01045; EIAV_GP45; 1.
DR Pfam: PF00971; EIAV_GP90; 1.
KW Coat protein; Glycoprotein; Polypeptide; Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 859
FT CHAIN 23 444
FT CHAIN 445 859
FT TRANSMEM 75 93
FT TRANSMEM 446 472
FT TRANSMEM 617 636
FT TRANSMEM 787 807
FT TRANSMEM 816 835
FT CARBOHYD 40 40
FT CARBOHYD 112 112
FT CARBOHYD 141 141
FT CARBOHYD 148 148
FT CARBOHYD 186 186
FT CARBOHYD 214 214
FT CARBOHYD 233 233
FT CARBOHYD 244 244
FT CARBOHYD 340 340
FT CARBOHYD 368 368
FT CARBOHYD 399 399
FT CARBOHYD 406 406
FT CARBOHYD 411 411
FT CARBOHYD 483 483
FT CARBOHYD 490 490
FT CARBOHYD 550 550
FT CARBOHYD 557 557
FT CARBOHYD 752 752
SQ SEQUENCE 859 AA: 97041 MW: 87180 PISC3A2B9 CRC64;
Query Match 73.9%; Score 34; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLQVQ 7
DB 265 RLQVQ 270
RESULT 12
OGAL_HUMAN STANDARD; PRT; 327 AA.
AC 095222;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE OLFACTORY RECEPTOR 6A1 (OLFACTORY RECEPTOR II-55) (ORII-55).
GN ORGAL.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99005533; PubMed-9787077;
RA Buettner J.A., Glusman G., Ben-Arie N., Ramos P., Lancet D.,
Evans G.A.;
RT "Organization and evolution of olfactory receptor genes on human
RT chromosome 11.";
RL Genomics 53:56-68(1998).
CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF065870; AAC70018.1; -.
DR InterPro: IPR000276; -.
DR InterPro: IPR000725; -.
DR Pfam: PF00001; 7em_1; 1.
DR PRINTS: PR00237; GPCR_HODOPSIN.
DR PRINTS: PR00245; OLFACTORYR.
DR PROSITE: PS00237; G_PROTEIN_REC_F1_1; FALSE_NEG.
DR PROSITE: PS00262; G_PROTEIN_REC_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Olfaction.
FT DOMAIN 1 26
FT TRANSMEM 27 50
FT DOMAIN 51 58
FT TRANSMEM 59 80
FT DOMAIN 81 105
FT TRANSMEM 106 125
FT DOMAIN 126 144
FT TRANSMEM 145 163
FT DOMAIN 164 201
FT TRANSMEM 202 224
FT TRANSMEM 225 241
FT TRANSMEM 242 264
FT DOMAIN 265 277
FT TRANSMEM 278 297
FT DOMAIN 298 320
FT DISULFID 102 194
FT CARBOHYD 5 5
FT CARBOHYD 191 191
SQ SEQUENCE 327 AA: 36049 MW: 69956A573ECFDA04 CRC64;
Query Match 71.7%; Score 33; DB 1; Length 327;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRLCVQ 6
DB 143 GRLCVQ 148
RESULT 13
WIF1_HUMAN STANDARD; PRT; 379 AA.
AC 0955W5;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE WNT INHIBITORY FACTOR 1 PRECURSOR (WIF-1).
GN WIF1.
OS Homo sapiens (Human).

Query Match	Best Local Similarity	Score 33;	DB 1;	Length 379;
Matches 6; Conservative	75.0%;	Pred. No. 30;		
Mismatches 0;		Mismatches 2;	Indels 0;	Gaps 0;
<p>Oy 1 GRLCVP08</p> <p>1 1 1 1 1 1</p> <p>Db 221 GGLCVTPG 228</p>				
<p>WT1_MOUSE</p> <p>WT1_MOUSE</p> <p>STANDARD;</p> <p>PRT; 379 AA.</p>				

AC	Q9WUJ1	09WUJ1	01-OCT-2000	(Rel. 40, Created)
DT	01-OCT-2000	(Rel. 40, Last sequence update)		
DT	01-OCT-2000	(Rel. 40, Last annotation update)		
DE	WNT INHIBITORY FACTOR 1 PRECURSOR (WIF-1).			
GN	WIF1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99215557; PubMed=10201374;			
RA	Shieh J.-C., Kodjabachian L., Rebhart M.L., Ratner A.,			
RT	Smallwood P.M., Samos C.H., Nusse R., David I.B., Nathans J.;			
RL	"A new secreted protein that binds to wnt proteins and inhibits their			
RT	activities.";			
RL	Nature 398:431-436(1999).			
CC	- FUNCTION: BINDS TO WNT PROTEINS AND INHIBITS THEIR ACTIVITIES. MAY			
CC	BE INVOLVED IN MESODERM SEGMENTATION.			
CC	- SUBCELLULAR LOCATION: SECRETED.			
CC	- TISSUE SPECIFICITY: EXPRESSION HIGHEST IN HEART AND LUNG. LOWER IN			
CC	BRAIN AND EYE.			
CC	- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its use			
CC	by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sdb.ch/announce/			
CC	or send an email to license@isb-sdb.ch).			
CC	-----			
DR	EMBL; AF122923; AAD25403.1; -			
DR	HSSP; P00740; IIXA.			
DR	MGD; MGI:1344332; Wt1f.			
DR	InterPro; IPR000561; -			
DR	InterPro; IPR002049; -			
DR	Pfam; PF000008; EGF_5.			
DR	PRINTS; PRO0011; EGYLAMININ.			
DR	PROSITE; PS00022; EGF_1; 5.			
DR	PROSITE; PS01186; EGF_2; 4.			
KW	Repeat; EGF-like domain; Signal; Developmental protein.			
KW	FT SIGNAL	1	28	POTENTIAL.
FT	CHAIN	29	379	
FT	DOMAIN	177	208	WNT INHIBITORY FACTOR 1.
FT	DOMAIN	209	240	EGF-LIKE 1.
FT	DOMAIN	241	272	EGF-LIKE 2.
FT	DOMAIN	273	304	EGF-LIKE 3.
FT	DOMAIN	305	336	EGF-LIKE 4.
FT	DISULFID	177	186	EGF-LIKE 5.
FT	DISULFID	182	192	POTENTIAL.
FT	DISULFID	198	200	POTENTIAL.
FT	DISULFID	209	218	POTENTIAL.
FT	DISULFID	214	224	POTENTIAL.
FT	DISULFID	230	232	POTENTIAL.
FT	DISULFID	241	250	POTENTIAL.
FT	DISULFID	246	256	POTENTIAL.
FT	DISULFID	262	264	POTENTIAL.
FT	DISULFID	273	282	POTENTIAL.
FT	DISULFID	278	288	POTENTIAL.
FT	DISULFID	294	296	POTENTIAL.
FT	DISULFID	305	314	POTENTIAL.
FT	DISULFID	310	320	POTENTIAL.
FT	DISULFID	326	328	POTENTIAL.
FT	CARBOHYD	88	88	POTENTIAL.
FT	CARBOHYD	245	245	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	245	245	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	379 AA;	41590 MW;	E3765F2642B2BC9A CRC64;

OY 1 GRLCVOPG 8
 1 1 1 1 1
 Db 221 GGLCVTPG 228

RESULT 15

GLPK_MYCPN STANDARD: PRT: 508 AA.
 ID GLPK_MYCPN
 AC P75064;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE)
 DE (GLYCEROKINASE) (GK).
 GN GLPK OR MPN050 OR MP104.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
 CC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID-2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 29342 / M129;
 RX MEDLINE-97105885; PubMed-8948633;
 RA Himmelfreich R., Hilbert H., Plagens H., Pirk E., Li B.-C.,
 Herrmann R.;
 RT "Complete genome analysis of the genome of the bacterium Mycoplasma
 pneumoniae.";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND
 METABOLISM.
 CC -1- CATALYTIC ACTIVITY: ATP + GLYCEROL = ADP + GLYCEROL 3-PHOSPHATE.
 CC -1- PATHWAY: RATE-LIMITING STEP IN GLYCEROL UTILIZATION.
 CC -1- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCONOKINASE /
 GLYCEROKINASE / XYLOKINASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AE000012; AAB95752.1; .
 DR HSSP: P08859; IGLB.
 DR InterPro: IPR000577; .
 DR Pfam: PF00370; FGGY_1.
 DR PROSITE: PS00445; FGGY_KINASES_2; 1.
 DR PROSITE: PS00933; FGGY_KINASES_1; 1.
 KW Glycerol metabolism; Transferase; Kinase; ATP-binding.
 FT NP_BIND 157 169 ATP (PROBABLE).
 SQ SEQUENCE 508 AA: 56591 MW: 86601332DFD62A6 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 508;
 Best Local Similarity 71.4%; Pred. NO. 39;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRLCVOP 7
 1 1 1 1 1
 Db 258 GGLCVTP 264

Search completed: June 13, 2001, 14:21:41
 Job time: 801 sec

This Page Blank (uspto)

THIS DOCUMENT IS UNCLASSIFIED

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:20:25 ; Search time 225.85 Seconds
(without alignments)
4.152 Million cell updates/sec

Title: PCT-US01-05825A-5
Perfect score: 46
Sequence: 1 GRICVOPG 8

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP unclassified:*
13: SP vertebrate:*
14: SP virus:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	87.0	447	2	086673 streptomyc
2	38	82.6	323	2	0918F5 0918f5 vibrio mimi
3	38	82.6	399	2	09R3V6 09r3v6 vibrio chol
4	38	82.6	399	2	09L706 09l7q6 vibrio chol
5	38	82.6	4123	4	075851 homo sapien
6	37	80.4	783	14	068952 hepatitis c
7	36	78.3	578	5	09W095 09w095 desophila
8	35	76.1	547	4	09S072 09s072 homo sapien
9	35	76.1	547	4	09W095 09w095 homo sapien
10	35	76.1	593	14	087093 087093 simian herp
11	34	73.9	88	5	09NLP5 09nlp5 leishmania
12	34	73.9	154	14	072771 072771 equine infe
13	34	73.9	154	14	072774 072774 equine infe
14	34	73.9	154	14	072775 072775 equine infe
15	34	73.9	154	14	072778 072778 equine infe
16	34	73.9	154	14	072779 072779 equine infe
17	34	73.9	154	14	072792 072792 equine infe
18	34	73.9	154	14	072793 072793 equine infe
19	34	73.9	154	14	072794 072794 equine infe

20	34	73.9	154	14	072795 equine infe
21	34	73.9	154	14	072796 equine infe
22	34	73.9	154	14	072797 equine infe
23	34	73.9	154	14	072798 equine infe
24	34	73.9	154	14	072801 equine infe
25	34	73.9	154	14	072807 equine infe
26	34	73.9	154	14	072812 equine infe
27	34	73.9	154	14	072813 equine infe
28	34	73.9	154	14	072814 equine infe
29	34	73.9	154	14	072819 equine infe
30	34	73.9	154	14	072800 equine infe
31	34	73.9	154	14	09WK02 09wk02 equine infe
32	34	73.9	154	14	09W9W0 09w9w0 equine infe
33	34	73.9	154	14	09W967 09w967 equine infe
34	34	73.9	160	14	072820 equine infe
35	34	73.9	161	14	072834 equine infe
36	34	73.9	162	14	072833 equine infe
37	34	73.9	163	14	070707 070707 equine infe
38	34	73.9	163	14	072772 equine infe
39	34	73.9	163	14	072773 equine infe
40	34	73.9	163	14	072788 equine infe
41	34	73.9	163	14	072789 equine infe
42	34	73.9	163	14	072790 equine infe
43	34	73.9	163	14	072791 equine infe
44	34	73.9	163	14	072809 equine infe
45	34	73.9	163	14	09W8F0 09w8f0 equine infe

ALIGNMENTS

RESULT 1
ID 086673 PRELIMINARY; PRT: 447 AA.
AC 086673;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE PUTATIVE LIPOPROTEIN.
GN SC4A2.1/C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID-1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Oliver K., Harris D.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE-97000351; PubMed-8843436;
RA Redenbach M., Kleser H.M., Denaplatte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.",
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL031182; CAA20169.1; -;
KW Lipoprotein.
SQ SEQUENCE 447 AA; 46712 MW; 809E0091B7834D80 CRC64;

Query Match 87.0%; Score 40; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GRICVOP 7

```

Db      31 GRLCVP 37

RESULT  2
ID      09L8F5      PRELIMINARY;      PRT;      323 AA.
AC      09L8F5;
DT      01-OCT-2000 (TREMBlrel. 15, Created)
DT      01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE      01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE      ZOT (FRAGMENT).
GN
OS      Vibrio mimicus.
OC      Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX      NCBI_Taxid=674;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=PT5;
RX      MEDLINE=20143766; PubMed=10678967;
RA      Boyd E.F., Moyer K.E., Shi L., Maldor M.K.;
RT      "Infectious CTXphi and the vibrio pathogenicity island prophage in
RT      Vibrio mimicus: evidence for recent horizontal transfer between V.
RL      mimicus and V. cholerae.";
DR      Infect. Immun. 68:1507-1513(2000).
FT      EMBL: AF207857; AAF40142.1; -;
FT      NON_TER      1      323
SQ      SEQUENCE      323 AA; 36306 MW; 01C12DAE9B873C3B CRC64;

Query Match
Best Local Similarity 82.6%; Score 38; DB 2; Length 323;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GRLCVP 8
Db      240 GRLCVP 247

RESULT  3
ID      09R3V6      PRELIMINARY;      PRT;      399 AA.
AC      09R3V6;
DT      01-MAY-2000 (TREMBlrel. 13, Created)
DT      01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT      01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE      ZONULAR OCCUDENS TOXIN (ZONA OCCUDENS TOXIN).
DE      ZOT OR VCI458.
GN      Vibrio cholerae.
OC      Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX      NCBI_Taxid=666;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=KNH002;
RA      Shin H.J., Park Y.C., Kim Y.C.;
RT      "Cloning and nucleotide sequence analysis of the virulence gene
RT      cassette from Vibrio cholerae KNH002 isolated in Korea.";
RL      Misalunhag Hoj1 35:205-210(1999).
RN      (2)
RP      SEQUENCE FROM N.A.
RC      STRAIN=0139-TOR OGAMA;
RA      Zhi-Tong H., Wei-Die Z., Xiang-Fu W.;
RT      "Cloning and Expression of zot Gene from Vibrio cholerae.";
RL      Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN      (3)
RP      SEQUENCE FROM N.A.
RC      STRAIN=EL TOR N16961 / SPROTYPE 01;
RX      MEDLINE=20406833; PubMed=10952301;
RA      Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Ginn M.L.,
RA      Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA      Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA      Esmolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,

```

```

RA      McDonald L., Uterback T., Fleisimann R.D., Niernan W.C., White O.,
RA      Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA      Fraser C.M.;
RT      "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT      cholerae.";
RL      Nature 406:477-483(2000).
DR      EMBL: AF175708; AAD51358.1; -;
DR      EMBL: AF123049; AAD26854.1; -;
DR      EMBL: AE004224; AAF94615.1; -;
DR      TIGR: VCI458; -;
SQ      SEQUENCE      399 AA; 44903 MW; 3C7424B758176774 CRC64;

Query Match
Best Local Similarity 82.6%; Score 38; DB 2; Length 399;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GRLCVP 8
Db      291 GRLCVP 298

RESULT  4
ID      09L706      PRELIMINARY;      PRT;      399 AA.
AC      09L706;
DT      01-OCT-2000 (TREMBlrel. 15, Created)
DT      01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE      01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE      ZOT.
GN      ZOT.
OC      Vibrio cholerae.
OC      Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX      NCBI_Taxid=666;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=86015;
RA      Kan B., Liu Y.Q., Qi G.M., Gao S.Y.;
RT      "Vibrio cholerae nct-CTXphi whole genome. Include rstr(Rstr),
RT      rsta(Rsta), rstrb(Rstrb), cep(Cep), orf(Orfu), ace(Ace) and zot(Zot)
RT      genes.";
RL      Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF220606; AAF29547.1; -;
SQ      SEQUENCE      399 AA; 44990 MW; CF6A3DBC9E23EE1 CRC64;

Query Match
Best Local Similarity 82.6%; Score 38; DB 2; Length 399;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GRLCVP 8
Db      291 GRLCVP 298

RESULT  5
ID      075851      PRELIMINARY;      PRT;      4123 AA.
AC      075851;
DT      01-NOV-1998 (TREMBlrel. 08, Created)
DT      01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT      01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE      WUGSC:H_D0751H13.1 PROTEIN (FRAGMENT).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_Taxid=9606;
RN      (1)
RP      SEQUENCE FROM N.A.
RA      Leonard S., Graves T., Strommatt C.;
RT      "The sequence of Homo sapiens PAC clone D0751H13.";
RL      Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN      (2)

```

RP SEQUENCE FROM N.A.
 RA Waterston R.H.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 EMBL: AC004877; AAC36301.1; -
 DR HSP: P01130; 1A7.
 DR INTERPRO: IPR000421; -
 DR INTERPRO: IPR000561; -
 DR INTERPRO: IPR000884; -
 DR INTERPRO: IPR000923; -
 DR INTERPRO: IPR001007; -
 DR INTERPRO: IPR001064; -
 DR INTERPRO: IPR001092; -
 DR INTERPRO: IPR001846; -
 DR INTERPRO: IPR002172; -
 DR INTERPRO: IPR002223; -
 DR INTERPRO: IPR002465; -
 DR INTERPRO: IPR002919; -
 DR PFAM: PF00057; 1d1_Recept_a; 11.
 DR PFAM: PF00090; tsp_1; 14.
 DR PFAM: PF00094; vwd; 5.
 DR PFAM: PF00754; F5_F8_type_C; 1.
 DR PFAM: PF01826; TTL; 9.
 DR PRINTS: PR00261; LDLRECEPTOR.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS00196; COPPER_BLUE; UNKNOWN_1.
 DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
 DR PROSITE: PS00280; BPTI_KUNITZ; UNKNOWN_1.
 DR PROSITE: PS00340; RECEPTOR_CYTOKINES_2; UNKNOWN_6.
 DR PROSITE: PS01209; LDLRA_1; 9.
 DR PROSITE: PS00068; LDLRA_2; 20.
 DR GLYCOPROTEIN.
 KW NON_TER
 SQ SEQUENCE 4123 AA; 434985 MW; 7AAB6F8DC0E012FB CRC64;

Query Match 82.6%; Score 38; DB 4; Length 4123;
 Best Local Similarity 75.0%; Pred. No. 99;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRUCVOPG 8
 DB 3234 GAICVOPG 3241

RESULT 6
 ID 068952 PRELIMINARY; PRT: 783 AA.
 AC 068952;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE (U4) CORE, E1, NS1/E2 AND NS2 GENES (FRAGMENTS).
 GN JKA.
 OS Hepatitis C virus.
 OC Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT 4:HEPATITIS C VIRUS, JKA;
 RX Honda M., Kaneko S., Unoura M., Kobayashi K., Murakami S.;
 HEPATITIS-93119270; PubMed=8380322;
 RL Arch. Virol. 128:163-169(1993).
 RT Isolated from 5 Japanese patients with hepatocellular carcinoma.";
 FT CHAIN 1 >191 CORE.
 FT NON_CONS 191 192
 FT CHAIN 192 >383 ENVELOPE 1.

FT NON_CONS 383 384
 FT CHAIN 384 >733 POTENTIAL.
 FT NON_CONS 733 734
 FT CHAIN 734 >783 NON-STRUCTURAL 2.
 FT NON_TER 783 783
 SQ SEQUENCE 783 AA; 85808 MW; A3145A3D310F9E5C CRC64;

Query Match 80.4%; Score 37; DB 14; Length 783;
 Best Local Similarity 75.0%; Pred. No. 36;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRUCVOPG 8
 DB 73 GRACVOPG 80

RESULT 7
 ID 09W095 PRELIMINARY; PRT: 578 AA.
 AC 09W095; 09W094;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE CG7995 PROTEIN.
 GN CG7995
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwa C.,
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Ruskern D.R., Pacleb J.M.,
 RA Palazozo M., Peltman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).

CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (AC Q9V6L1) AND
 CC SHORT ISOFORM (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
 DR EMBL: AE003472; AAF47558.1; -
 DR EMBL: AE003472; AAF47559.1; -
 DR HSSP: P08859; IGIC.
 DR FLVASE: P890035266; CG7995.
 DR INTERPRO: IPR000577; -
 DR PFM: PF00370; FGGY; 1.
 KW Alternative splicing; Hypothetical protein.
 FT VARSPLIC 1 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 578 AA; 64905 MW; D77429D03B658D3 CRC64.

Query Match
 Best Local Similarity 78.3%; Score 35; DB 5; Length 578;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRICVOPG 8
 DB 286 GOMCVKPG 293

RESULT 8
 ID 095072 PRELIMINARY; PRT: 547 AA.
 AC 095072;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE HHR21SPB (FRAGMENT).
 GN HHR21SPB.
 OS Homo sapiens (Human).
 CC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RA McKay M.J., van der Spek P., Kanaar R., Bootsma D., Hoelmakers J.H.;
 RT "hHR21SPB, a second human gene homologous to the rad21
 RT Schizosaccharomyces pombe DNA double strand break repair gene."
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF006284; AAD01193.1; -
 FT NON_TER 547
 SQ SEQUENCE 547 AA; 62613 MW; 544AA31F2D038D7B CRC64;

Query Match
 Best Local Similarity 76.1%; Score 35; DB 4; Length 547;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRICVOPG 8
 DB 536 GRLLIOPG 543

RESULT 9
 ID 09NV09 PRELIMINARY; PRT: 547 AA.
 AC 09NV09;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE CDNA FLJ10573 FIS, CLONE NT2RP2003177.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara K., Takeuchi K.,

RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y., Oshima A.;
 RT "NEDO human cDNA sequencing project."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK001435; BAA91690.1; -
 SQ SEQUENCE 547 AA; 62585 MW; 05A1CFC8746C3BB CRC64;

Query Match
 Best Local Similarity 76.1%; Score 35; DB 4; Length 547;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRICVOPG 8
 DB 536 GRLLIOPG 543

RESULT 10
 ID 087093 PRELIMINARY; PRT: 593 AA.
 AC 087093;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE GLYCOPROTEIN G (HOMOLOGUE OF HSV-2 us4).
 OS Simian herpes B virus (Cercopithecidae herpesvirus 1) (shbv).
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=10325;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CYNO 2;
 RA MEDLINE=96005037; PubMed=7561753;
 RA Slomka M.J., Harrington L., Arnold C., Norcott J.P.N., Brown D.W.;
 RT "Complete nucleotide sequence of the herpesvirus simiae glycoprotein G
 RT gene and its expression as an immunogenic fusion protein in
 RL bacteria."
 RL J. Gen. Virol. 76:2161-2168(1995).
 DR EMBL: Z46268; CAA86431.1; -
 SQ SEQUENCE 593 AA; 60731 MW; 79BD2686FA82B8B5 CRC64;

Query Match
 Best Local Similarity 76.1%; Score 35; DB 14; Length 593;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVOPG 8
 DB 248 GRICVOPG 255

RESULT 11
 ID 09NLP5 PRELIMINARY; PRT: 88 AA.
 AC 09NLP5;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE HYPOTHEICAL 9.2 KDA PROTEIN (FRAGMENT).
 GN LM26.397.
 OS Leishmania major.
 CC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIDELIN;
 RA Murphy L., Quail M., Harris D., Rajandream M., Irens A., Barrell B.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL160493; CAB88015.1; -
 KW Hypothetical protein.
 FT NON_TER 88
 SQ SEQUENCE 88 AA; 9238 MW; 717B17CABD8C031A CRC64;

Query Match 73.9%; Score 34; DB 5; Length 88;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVOP 7
|11111
DB 56 RLCVOP 61

RESULT 12

072771 PRELIMINARY; PRT; 154 AA.
AC 072771:
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
GN ENV.
OS Equine infectious anemia virus.
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-V70:
RA Zheng Y., Nakaya T., Sentsui H., Kameoka M., Kishi M., Takahashi H.,
RA Kono Y., Ikuta K.;
RL J. Gen. Virol. 0:0-0(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-V70:
RX MEDLINE=97333334; PubMed=9188568;
RA Zheng Y.-H., Sentsui H., Nakaya T., Kono Y., Ikuta K.;
RT "In vivo dynamics of equine infectious anemia viruses emerging during
RT febrile episodes: Insertions/duplications at the principal
RT neutralizing domain."
RL J. Virol. 71:5031-5039(1997).
DR EMBL: D87721; BAA28364.1; -.
DR INTERPRO: IPR001361; -.
DR PFAM: PF00971; EIAV_GP90; 1.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 154
SQ SEQUENCE 154 AA; 17777 MW; 292A3C53462E67EB CRC64;

Query Match 73.9%; Score 34; DB 14; Length 154;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVOP 7
|11111
DB 76 RLCVOP 81

RESULT 13

072774 PRELIMINARY; PRT; 154 AA.
AC 072774:
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
GN ENV.
OS Equine infectious anemia virus.
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-F2V;
RA Zheng Y., Nakaya T., Sentsui H., Kameoka M., Kishi M., Takahashi H.,
RA Kono Y., Ikuta K.;
RL J. Gen. Virol. 0:0-0(1996).

RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN-F2V:
RX MEDLINE=97333334; PubMed=9188568;
RA Zheng Y.-H., Sentsui H., Nakaya T., Kono Y., Ikuta K.;
RT "In vivo dynamics of equine infectious anemia viruses emerging during
RT febrile episodes: Insertions/duplications at the principal
RT neutralizing domain."
RL J. Virol. 71:5031-5039(1997).
DR EMBL: D87724; BAA28366.1; -.
DR INTERPRO: IPR001361; -.
DR PFAM: PF00971; EIAV_GP90; 1.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 154
SQ SEQUENCE 154 AA; 17803 MW; C2569D3299A4632C CRC64;

Query Match 73.9%; Score 34; DB 14; Length 154;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVOP 7
|11111
DB 76 RLCVOP 81

RESULT 14

072775 PRELIMINARY; PRT; 154 AA.
AC 072775:
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
GN ENV.
OS Equine infectious anemia virus.
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-F2V:
RA Zheng Y., Nakaya T., Sentsui H., Kameoka M., Kishi M., Takahashi H.,
RA Kono Y., Ikuta K.;
RL J. Gen. Virol. 0:0-0(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-F2V:
RX MEDLINE=97333334; PubMed=9188568;
RA Zheng Y.-H., Sentsui H., Nakaya T., Kono Y., Ikuta K.;
RT "In vivo dynamics of equine infectious anemia viruses emerging during
RT febrile episodes: Insertions/duplications at the principal
RT neutralizing domain."
RL J. Virol. 71:5031-5039(1997).
DR EMBL: D87725; BAA28367.1; -.
DR INTERPRO: IPR001361; -.
DR PFAM: PF00971; EIAV_GP90; 1.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 154
SQ SEQUENCE 154 AA; 17759 MW; 38EA5B4BD5B5C34 CRC64;

Query Match 73.9%; Score 34; DB 14; Length 154;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVOP 7
|11111
DB 76 RLCVOP 81

RESULT 15

```

072778
ID 072778 PRELIMINARY; PRT; 154 AA.
AC 072778;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
GN ENV.
OS Equine infectious anemia virus.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F4V;
RA Zheng Y., Nakaya T., Sentsui H., Kameoka M., Kishi M., Takahashi H.,
RL J. Gen. Virol. 0:0-0(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=F4V;
RX MEDLINE=9733334; PubMed=9188568;
RA Zheng Y.-H., Sentsui H., Nakaya T., Kono Y., Ikuta K.;
RT "In vivo dynamics of equine infectious anemia viruses emerging during
RT febrile episodes: insertions/duplications at the principal
RT neutralizing domain.";
RL J. Virol. 71:5031-5039(1997).
DR EMBL; D87728; BAA28370.1; -.
DR INTERPRO: IPR001361; -.
DR PFM: PFO0971; EIAV_GP90; 1.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 154
FT NON_TER 154
SQ SEQUENCE 154 AA; 17706 MW; 393282E5A70DC7EB CRC64;

```

```

Query Match 73.98; Score 34; DB 14; Length 154;
Best Local Similarity 100.08; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 RLCVOP 7
   |||||
Db 76 RLCVOP 81

```

Search completed: June 13, 2001, 14:20:26
Job time: 726 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:16:32 ; Search time 118.55 Seconds
(without alignments)
1.296 Million cell updates/sec

Title: PCT-US01-05825a-5
Perfect score: 46
Sequence: 1 GRICVPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCVUS.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	82.6	400	2 US-08-624-601-8	Sequence 8, Appli
2	35	76.1	76	3 US-08-851-843A-205	Sequence 205, App
3	35	76.1	76	4 US-08-974-549A-324	Sequence 324, App
4	34	73.9	130	4 US-09-347-819-8	Sequence 8, Appli
5	31	67.4	92	2 US-07-728-215-39	Sequence 39, Appli
6	31	67.4	92	2 US-07-728-215-41	Sequence 41, Appli
7	31	67.4	118	2 US-08-652-816A-12	Sequence 12, Appli
8	31	67.4	471	5 PCT-US95-07820-2	Sequence 2, Appli
9	31	67.4	471	5 PCT-US95-07820-2	Sequence 2, Appli
10	31	67.4	554	1 US-08-445-586-7	Sequence 7, Appli
11	31	67.4	556	1 US-08-445-586-2	Sequence 2, Appli
12	31	67.4	567	3 US-09-188-811-2	Sequence 2, Appli
13	31	67.4	718	1 US-08-444-792-4	Sequence 4, Appli
14	31	67.4	718	1 US-08-445-042-4	Sequence 4, Appli
15	31	67.4	788	1 US-07-728-215-32	Sequence 32, Appli
16	31	67.4	907	1 US-08-349-006-2	Sequence 2, Appli
17	31	67.4	907	3 US-08-804-439A-19	Sequence 19, Appli
18	31	67.4	907	3 US-08-720-229-19	Sequence 19, Appli
19	31	67.4	907	5 PCT-US94-04180-2	Sequence 2, Appli
20	30	65.2	190	1 US-08-106-981-2	Sequence 2, Appli
21	30	65.2	257	1 US-07-828-790B-2	Sequence 2, Appli
22	30	65.2	266	4 US-09-347-819-10	Sequence 10, Appli
23	30	65.2	297	2 US-09-006-535-4	Sequence 4, Appli
24	30	65.2	328	1 US-08-118-270-42	Sequence 42, Appli
25	30	65.2	328	5 PCT-US93-08528-42	Sequence 42, Appli
26	30	65.2	384	1 US-07-937-609-17	Sequence 17, Appli
27	30	65.2	384	4 US-08-029-170-17	Sequence 17, Appli

28	30	65.2	384	5 PCT-US92-02091-2	Sequence 2, Appli
29	30	65.2	384	5 PCT-US92-02091-4	Sequence 4, Appli
30	30	65.2	397	2 US-08-989-478-14	Sequence 14, Appli
31	30	65.2	397	3 US-08-996-685-14	Sequence 14, Appli
32	30	65.2	411	3 US-08-964-652-2	Sequence 2, Appli
33	30	65.2	469	2 US-08-989-478-10	Sequence 10, Appli
34	30	65.2	469	3 US-08-996-685-10	Sequence 10, Appli
35	30	65.2	521	2 US-08-989-478-12	Sequence 12, Appli
36	30	65.2	521	3 US-08-996-685-12	Sequence 12, Appli
37	30	65.2	533	4 US-09-347-819-2	Sequence 2, Appli
38	30	65.2	574	4 US-09-062-440-2	Sequence 2, Appli
39	30	65.2	574	4 US-09-062-440-4	Sequence 4, Appli
40	30	65.2	574	4 US-09-062-440-6	Sequence 6, Appli
41	30	65.2	593	2 US-08-989-478-2	Sequence 2, Appli
42	30	65.2	593	3 US-08-989-478-8	Sequence 8, Appli
43	30	65.2	593	3 US-08-996-685-8	Sequence 8, Appli
44	30	65.2	593	3 US-08-996-685-8	Sequence 8, Appli
45	30	65.2	593	3 US-08-880-179-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-624-601-8
Sequence 8, Application US/08624601
Patent No. 5882653
GENERAL INFORMATION:
APPLICANT: Kaper Dr., James B.
TITLE OF INVENTION: Vibrio cholerae O1 (CVD111) and non-O1
TITLE OF INVENTION: (CVD112 and CVD112RM) serogroup vaccine strains, methods
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESS: Spencer & Frank
STREET: 1100 New York Ave. N.W. Suite 300 East
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/624,601
FILING DATE: 08-APR-1996
CLASSIFICATION: A24
ATTORNEY/AGENT INFORMATION:
NAME: Schellier Dr., John W.
REGISTRATION NUMBER: 26,031
REFERENCE/DOCKET NUMBER: BANCZ0019P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)414-4000
TELEFAX: (202)414-4040
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Vibrio cholerae
STRAIN: El Tor 7946
IMMEDIATE SOURCE:
CLONE: zot
US-08-624-601-8

Query Match 82.6%; Score 38; DB 2; Length 400;
Best Local Similarity 87.5%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRLCVOPG 8
DB 291 GRLCVODG 298

RESULT 2

US-08-851-843A-205
Sequence 205, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-0029300S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 205:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-851-843A-205

Query Match 76.1%; Score 35; DB 3; Length 76;
Best Local Similarity 75.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRLCVOPG 8
DB 49 GRLCAGPG 56

RESULT 3

US-08-974-549A-324
Sequence 324, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-0026100S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 324:

SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-324

Query Match 76.1%; Score 35; DB 4; Length 76;
Best Local Similarity 75.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRLCVOPG 8
|||
DB 49 GRLCAGPG 56

RESULT 4
US-09-347-819-8
Sequence 8, Application US/09347819
Patent No. 6184036
GENERAL INFORMATION:
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Rafalski, J. Antoni
APPLICANT: Falco, S. Carl
TITLE OF INVENTION: Ornithine Biosynthesis Enzymes
FILE REFERENCE: BB-1174-C
CURRENT APPLICATION NUMBER: US/09/347,819
CURRENT FILING DATE: 1999-07-02
EARLIER APPLICATION NUMBER: 60/093,209
EARLIER FILING DATE: July 17, 1998
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Microsoft Office 97
SEQ ID NO 8
LENGTH: 130
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: UNSURE
LOCATION: (109)
US-09-347-819-8

Query Match 73.9%; Score 34; DB 4; Length 130;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LCVOPG 8
|||
DB 74 LCVOPG 79

RESULT 5
US-07-728-215-39
Sequence 39, Application US/07728215
Patent No. 5962643
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Pytela, Robert
TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/728,215
FILING DATE: 19910711
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P31 8717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-728-215-39

Query Match 67.4%; Score 31; DB 2; Length 92;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 LCVOPG 8
|||
DB 82 VCIOPG 87

RESULT 6
US-07-728-215-41
Sequence 41, Application US/07728215
Patent No. 5962643
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Pytela, Robert
TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/728,215
FILING DATE: 19910711
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P31 8717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-728-215-41

Query Match 67.4%; Score 31; DB 2; Length 92;
 Best Local Similarity 66.7%; Pred. No. 61;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 GRLVOPG 8
 DB 82 VCIOPG 87

RESULT 7

US-08-652-816A-12
 ; Sequence 12, Application US/08652816A
 ; Patent No. 5872215
 ; GENERAL INFORMATION:
 ; APPLICANT: Osbourn, JK
 ; APPLICANT: Allen, DJ
 ; TITLE OF INVENTION: Specific binding members, materials and
 ; TITLE OF INVENTION: methods.
 ; NUMBER OF SEQUENCES: 53
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/652,816A
 ; FILING DATE: 23-MAY-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9125579.4
 ; FILING DATE: 02-DEC-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9125579.8
 ; FILING DATE: 02-DEC-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9206318.9
 ; FILING DATE: 24-MAR-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9206372.6
 ; FILING DATE: 23-SEP-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9525004.9
 ; FILING DATE: 07-DEC-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9610824.6
 ; FILING DATE: 23-MAY-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/GB92/02240
 ; FILING DATE: 02-DEC-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/244,597
 ; FILING DATE: 01-JUN-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: David W. Clough
 ; REGISTRATION NUMBER: 36,107
 ; REFERENCE/DOCKET NUMBER: 28111/33308
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-474-6300
 ; INFORMATION FOR SEQ. ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 118 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-08-652-816A-12

Query Match 67.4%; Score 31; DB 2; Length 118;
 Best Local Similarity 75.0%; Pred. No. 78;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRLVOPG 8
 DB 8 GRLVOPG 15

RESULT 8

US-08-466-583-2
 ; Sequence 2, Application US/08466583
 ; Patent No. 5919998
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandurski, Robert S.
 ; APPLICANT: Szczygielowski, Krzysztof
 ; TITLE OF INVENTION: Genetic Control of Plant Hormone Levels
 ; TITLE OF INVENTION: and Plant Growth.
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee & Winer, P.C.
 ; STREET: 5370 Manhattan Circle, Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: United States of America
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/466,583
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/265,427
 ; FILING DATE: 24-JUN-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ferber, Donna M.
 ; REGISTRATION NUMBER: 33,878
 ; REFERENCE/DOCKET NUMBER: 11-94A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 303/499-8080
 ; TELEFAX: 303/499-8089
 ; TELEX: 49617824
 ; INFORMATION FOR SEQ. ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 471 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-466-583-2

Query Match 67.4%; Score 31; DB 2; Length 471;
 Best Local Similarity 75.0%; Pred. No. 3e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRLVOPG 8
 DB 144 GRLVOPG 151

RESULT 9

PCT-US95-07820-2
 ; Sequence 2, Application PC/TUS9507820
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandurski, Robert S.
 ; APPLICANT: Szczygielowski, Krzysztof

TITLE OF INVENTION: Genetic Control of Plant Hormone Levels
TITLE OF INVENTION: and Plant Growth.
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee & Winer, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: United States of America
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07820
FILING DATE: 19-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,427
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Farber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 11-94B PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
TELEX: 49617824
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 471 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-07820-2

Query Match 67.4%; Score 31; DB 5; Length 471;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVOPG 8
DB 144 GRILVPPG 151

RESULT 10
US-08-445-586-7
Sequence 7, Application US/08445586
Patent No. 5627050
GENERAL INFORMATION:
APPLICANT: Takeshita, Sunao
APPLICANT: Ito, Toshimi
APPLICANT: Otawara-Hamamoto, Yoko
APPLICANT: Amann, Egon
TITLE OF INVENTION: Bone-Related Sulfatase-Like Protein and
TITLE OF INVENTION: Process for Its Production
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,586
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/111,887
FILING DATE: 26-AUG-1993
APPLICATION NUMBER: JP 230030/92
FILING DATE: 28-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 324034/92
FILING DATE: 03-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Forman, David S.
REGISTRATION NUMBER: 33,694
REFERENCE/DOCKET NUMBER: 02481.1322-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 554 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-445-586-7

Query Match 67.4%; Score 31; DB 1; Length 554;
Best Local Similarity 75.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVOPG 8
DB 47 GRILVPPG 54

RESULT 11
US-08-445-586-2
Sequence 2, Application US/08445586
Patent No. 5627050
GENERAL INFORMATION:
APPLICANT: Takeshita, Sunao
APPLICANT: Ito, Toshimi
APPLICANT: Otawara-Hamamoto, Yoko
APPLICANT: Amann, Egon
TITLE OF INVENTION: Bone-Related Sulfatase-Like Protein and
TITLE OF INVENTION: Process for Its Production
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,586
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/111,887
FILING DATE: 26-AUG-1993
APPLICATION NUMBER: JP 230030/92
FILING DATE: 28-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 324034/92

FILING DATE: 03-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Forman, David S.
REGISTRATION NUMBER: 33,694
REFERENCE/DOCKET NUMBER: 02481.1322-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 556 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-445-586-2

Query Match 67.4%; Score 31; DB 1; Length 556;
Best Local Similarity 75.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRLCOPG 8
DB 47 GRLCOPG 54

RESULT 12
US-09-188-811-2
Sequence 2, Application US/09188811
Patent No. 6037148
GENERAL INFORMATION:
APPLICANT: Khodadoust, Mehman
TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: MNI-046CP
CURRENT APPLICATION NUMBER: US/09/188,811
CURRENT FILING DATE: 1998-11-09
EARLIER APPLICATION NUMBER: 09/163,116
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 2
LENGTH: 567
TYPE: PRT
ORGANISM: Homo sapiens
US-09-188-811-2

Query Match 67.4%; Score 31; DB 3; Length 567;
Best Local Similarity 71.4%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 RLCVOPG 8
DB 59 RLCVOPG 65

RESULT 13
US-08-444-792-4
Sequence 4, Application US/08444792
Patent No. 5726037
GENERAL INFORMATION:
APPLICANT: Bodary, Sarah C.
APPLICANT: Gorman, Cornelia M.
APPLICANT: McLean, John W.
APPLICANT: Napier, Mary A.
TITLE OF INVENTION: METHOD FOR PREPARING WATER SOLUBLE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco

STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpath (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,792
FILING DATE: 19-May-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380227
FILING DATE: 30-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/218878
FILING DATE: 28-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/821337
FILING DATE: 13-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/444490
FILING DATE: 01-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/290224
FILING DATE: 22-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: P0552P1C3D4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/952-9881
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 718 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-444-792-4

Query Match 67.4%; Score 31; DB 1; Length 718;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 LCVOPG 8
DB 613 VCIOPG 618

RESULT 14
US-08-445-042-4
Sequence 4, Application US/08445042
Patent No. 5726290
GENERAL INFORMATION:
APPLICANT: Bodary, Sarah C.
APPLICANT: Gorman, Cornelia M.
APPLICANT: McLean, John W.
APPLICANT: Napier, Mary A.
TITLE OF INVENTION: METHOD FOR PREPARING WATER SOLUBLE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk


```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatlin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,042
FILING DATE: 19-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380227
FILING DATE: 30-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/218878
FILING DATE: 28-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/821337
FILING DATE: 13-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/444490
FILING DATE: 01-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/290224
FILING DATE: 22-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: P0552PIC3D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 718 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-445-042-4

```

```

Query Match      67.4% Score 31: DB 1: Length 718;
Best Local Similarity 66.7% Pred. No. 4.5e+02;
Matches 4: Conservative 2: Mismatches 0: Indels 0: Gaps 0:
OY 3 LCVOPG 8
:1:111
DB 613 VCIOPG 618

```

```

RESULT 15
US-07-728-215-32
Sequence 32, Application US/07728215
Patent No. 5962643
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Pytela, Robert
TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
thereof
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Preity, Schroeder, Brueggemann & Clark
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/728,215
FILING DATE: 19910711

```

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P31 8717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 788 amino acids
TYPE: AMINO ACID
STRANDEDNESS: Single
TOPOLOGY: Linear
US-07-728-215-32

```

```

Query Match      67.4% Score 31: DB 2: Length 788;
Best Local Similarity 66.7% Pred. No. 4.9e+02;
Matches 4: Conservative 2: Mismatches 0: Indels 0: Gaps 0:
OY 3 LCVOPG 8
:1:111
DB 613 VCIOPG 618

```

```

Search completed: June 13, 2001, 14:16:32
Job time: 493 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:14:32 ; Search time 229.28 Seconds
(without alignments)
1.995 Million cell updates/sec

Title: PCT-US01-05825A-6

Perfect score: 45

Sequence: 1 GRICVDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401:*

- 1: /SID6/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SID6/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SID6/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SID6/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SID6/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SID6/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SID6/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SID6/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SID6/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SID6/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SID6/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SID6/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SID6/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SID6/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SID6/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SID6/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SID6/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SID6/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SID6/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SID6/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SID6/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SID6/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	8	21	Peptide antagonist
2	45	100.0	399	13	Zonula occludens t
3	42	93.3	8	21	Peptide antagonist
4	38	84.4	8	21	Peptide antagonist
5	38	84.4	8	21	Peptide antagonist
6	37	82.2	8	21	Peptide antagonist
7	36	80.0	742	16	Human zona pelluci
8	36	80.0	745	15	Human zona pelluci
9	36	80.0	745	20	Human zona pelluci
10	36	80.0	745	20	Human zpa protein.
11	36	80.0	745	21	Human zona pelluci

12	36	80.0	745	21	Y52689	Human oocyte zona
13	36	80.0	745	21	Y52180	Human zona pelluci
14	36	80.0	745	21	Y52984	Human zona pelluci
15	35	77.8	8	21	Y79105	Peptide antagonist
16	35	77.8	8	21	Y79112	Peptide antagonist
17	35	77.8	8	21	Y79118	Peptide antagonist
18	34	75.6	713	15	R60101	Peptide antagonist
19	34	75.6	713	15	R55198	Canine zona pelluc
20	34	75.6	715	15	R55198	Canine zona pelluc
21	34	75.6	715	20	W81808	Canine zona pelluc
22	34	75.6	715	20	W81808	Canine zona pelluc
23	34	75.6	715	21	Y52680	Canine oocyte zona
24	34	75.6	715	21	Y52171	Canine oocyte zona
25	34	75.6	715	21	Y52975	Canine zona pelluc
26	34	75.6	716	15	R55200	Feline zona pelluc
27	34	75.6	716	15	R55200	Feline zona pelluc
28	34	75.6	716	20	W42473	Feline zona pelluc
29	34	75.6	716	20	W81810	Feline zpa protein
30	34	75.6	716	21	Y83208	Feline zona pelluc
31	34	75.6	716	21	Y52682	Feline oocyte zona
32	34	75.6	716	21	Y52173	Feline zona pelluc
33	34	75.6	716	21	Y52977	Feline zona pelluc
34	34	75.6	716	21	W63748	Human semaphorin.
35	34	75.6	775	19	W63748	Mouse semaphorin H
36	34	75.6	775	20	Y43090	Mouse semaphorin H
37	33	73.3	196	20	Y36831	Mouse semaphorin H
38	32	71.1	8	21	Y79108	Amino acid sequenc
39	32	71.1	115	17	R91223	Peptide antagonist
40	32	71.1	149	21	B25668	Mouse pancreas G-p
41	32	71.1	328	17	R91224	Human secreted pro
42	32	71.1	328	17	R91225	Mouse pancreas G-p
43	32	71.1	328	18	W09433	Human placenta G-p
44	31	68.9	8	21	Y79121	Human placenta pur
45	31	68.9	190	16	R72501	Peptide antagonist
						Wheat cold toleran

ALIGNMENTS

RESULT	ID	Y79110 standard; Peptide: 8 AA.	ALIGNMENTS
XX	XX	Y79110;	
XX	XX	05-JUN-2000 (first entry)	
DE	DE	Peptide antagonist of zonulin.	
KW	KW	Zonulin; antagonist; zonula occludens toxin receptor;	
KW	KW	blood-brain barrier; antiinflammatory; cerebroprotective;	
KW	KW	neuroprotective; dermatological; antidiabetic; antiviral;	
KW	KW	antibacterial; cytostatic; anti-HIV; vulnary; antiallergic;	
KW	KW	hypotensive; immunosuppressive; antiparasitic; vasotropic;	
KW	KW	gastrointestinal inflammation; therapy.	
OS	OS	Synthetic.	
XX	XX	WO200007609-A1.	
XX	XX	17-FEB-2000.	
XX	XX	28-JUL-1999; 99WO-US16683.	
PR	PR	03-AUG-1998; 98US-0127815.	
PA	PA	(UYMA-) UNIV MARYLAND BALTIMORE.	
XX	XX	Fasano A;	
XX	XX	PI	
XX	XX	WPI; 2000-205565/18.	
DR	DR	New peptide antagonist of zonulin useful as antiinflammatory agent for	
PT	PT		

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PS shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 Claim 1: Page 42; 69pp; English.
 PS
 XX

This present sequence is that of a peptide antagonist of zonulin (2), one of 25 such peptides (see Y79105-29) of the invention, which bind to a zonula occludens toxin (ZOT) receptor, yet do not physiologically modulate the opening of mammalian tight junctions (TJ). The peptide antagonists are based on a common motif of ZOT and human zonulins, which is believed to be critical for receptor binding. They can be prepared by chemical synthesis or by use of recombinant DNA techniques. The peptide antagonists are used as an antiinflammatory agents in the treatment of gastrointestinal inflammation, where they bind to the ZOT receptor in the intestine and yet does not physiologically modulate the opening of TJ in the intestine. Gastrointestinal inflammation conditions give rise to increased intestinal permeability and the peptide is useful for treating intestinal conditions that cause protein losing enteropathy caused by infection, e.g. Clostridium difficile infection, enterocolitis, shigellosis, viral gastroenteritis, parasite infestation, bacterial overgrowth, whipple's disease, diseases with mucosal erosion or ulcerations, e.g. gastritis, gastric cancer, collagenous colitis, inflammatory bowel disease, diseases marked by lymphatic obstruction, e.g. congenital intestinal lymphangiectasia, sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical correction of congenital heart disease with Fontan's operation, mucosal diseases without ulceration, e.g. Menetrier's disease, coeliac disease, eosinophilic gastroenteritis, and immune diseases, e.g. systemic lupus erythematosus or food allergies, primarily to milk.

50 Sequence 8 AA;

Query Match	100.0%;	Score 45;	DB 21;	Length 8;
Best Local Similarity	100.0%;	Pred. No. 3.2e+05;		
Matches	8; Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY	1	GRLCVQDG	8
Db	1	grlcvgdq	8

RESULT 2

ID	R20006	standard; Protein; 399 AA.
----	--------	----------------------------

AC R20006;

DT 31-MAR-1992 (first entry)

Zonula occludens toxin.

20T; cholera; vaccine; enterotoxin; diarrhoea.

OS Vibrio cholerae.
VY

PN W09118979-A.
XX

PD 12-DEC-1991.
xy

PF 05-JUN-1991; 91WO-US03812.

PR 05-JUN-1990; 90US-05333315.

PA (UYMA-) UNIV MARYLAND BALTI.

PI Kaper JB, Baudry-Maurelli B, Fasano A;

DR WPI; 1992-007465/01.

XX

PT New Vibrio cholerae strains - comprise restriction endonuclease
PT fragment encoding toxin, used as vaccines against cholera
XX
PS Disclosure: Fig 18; 83pp; English.
VV

The amino acid sequence is that of the zonula occludens toxin (ZOT). It may be responsible for diarrhoea in some strains of cholera and the ZOT gene or fragments of it are deleted from strains of *Vibrio cholerae* (V.c.) to be used as vaccines. These V.c. strains have 100% efficiency in protecting humans against subsequent infection with a strain of a similar serotype and avoid undesirable side effects such as diarrhoea, nausea and cramping. Cultures of these strains may be used for prodn. of vaccines against cholera.

50 Sequence 399 AA;

Query Match	100.0%	Score 45	DB 13	Length 399
Best Local	Similarity 100.0%	Pred. NO. 1.4		
Matches 8	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```
QY      1  GRLCVQDG  8
          |||||
Db .    291  grlcvgdq 298
```

RESULT 3

ID	Y79106	standard; Peptide; 8 AA.
...		

AC Y79106;

DT 05-JUN-2000 (first entry)

Peptide antagonist of zonulin

KM zonulin; antagonists; zonula occludens toxin receptor;
KM blood-brain barrier; antiinflammatory; cerebroprotective;
KM neuroprotective; dermatological; antitumor; antiviral;
KM antibacterial; cytostatic; anti-HIV; vulnereary; antiallergic
KM hypotensive; immunosuppressive; antiparasitic; vasotropic;
KM gastrointestinal inflammation; therapy.

OS Synthetic.

PN WO200007609-A1.

PD 17-FEB-2000

PF : 28-JUL-1999; 99WO-US16683.

PR 03-AUG-1998; 98US-0127815.

PA (UYMA-) UNIV MARYLAND BALTIMORE

PI Fasano A;

DR WPI; 2000-205565/18.

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
XX
PS Claim 1, Page 41, 69pp; English.
XX

This present sequence is that of a peptide antagonist of zonulin (2), one of 25 such peptides (see Y79105-29) of the invention, CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not CC physiologically modulate the opening of mammalian tight junctions CC (TJ). The peptide antagonists are based on a common motif of ZOT CC and human zonulins, which is believed to be critical for receptor CC binding. They can be prepared by chemical syntheses or by use of CC recombinant DNA techniques. The peptide antagonists are used as an

CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.

SO Sequence 8 AA;

Query Match 93.3%; Score 42; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. No. 3.2e+05;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRUCVQDG 8
11:111111
Db 1 grvcvqdg 8

RESULT 4
Y79109
ID Y79109 standard; Peptide: 8 AA.

XX Y79109;

DT 05-JUN-2000 (first entry)

XX Peptide antagonist of zonulin.

XX Zonulin; antagonist; zonula occludens toxin receptor;
XX blood-brain barrier; antiinflammatory; cerebroprotective;
XX neuroprotective; dermatological; anticancer; antiviral;
XX antibacterial; cytostatic; anti-HIV; vulnerrary; antiallergic;
XX hypotensive; immunosuppressive; antiparasitic; vasotropic;
XX gastrointestinal inflammation; therapy.

OS Synthetic.

PN WO200007609-A1.

PD 17-FEB-2000.

XX 28-JUL-1999; 99WO-US16683.

XX 03-AUG-1998; 98US-0127815.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

PA Fasano A;

PI WPI; 2000-205565/18.

XX New peptide antagonist of zonulin useful as antiinflammatory agent for
XX treating cerebral ischemia, stroke, cerebral edema, gastritis,
XX shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
XX Claim 1; Page 42; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not

CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.

SO Sequence 8 AA;

Query Match 84.4%; Score 38; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. No. 3.2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRUCVQDG 8
11111111
Db 1 grlvcvpg 8

RESULT 5
Y79122
ID Y79122 standard; Peptide: 8 AA.

XX Y79122;

DT 05-JUN-2000 (first entry)

XX Peptide antagonist of zonulin.

XX Zonulin; antagonist; zonula occludens toxin receptor;
XX blood-brain barrier; antiinflammatory; cerebroprotective;
XX neuroprotective; dermatological; anticancer; antiviral;
XX antibacterial; cytostatic; anti-HIV; vulnerrary; antiallergic;
XX hypotensive; immunosuppressive; antiparasitic; vasotropic;
XX gastrointestinal inflammation; therapy.

OS Synthetic.

PN WO200007609-A1.

PD 17-FEB-2000.

XX 28-JUL-1999; 99WO-US16683.

XX 03-AUG-1998; 98US-0127815.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

PA Fasano A;

PI WPI; 2000-205565/18.

XX New peptide antagonist of zonulin useful as antiinflammatory agent for
XX treating cerebral ischemia, stroke, cerebral edema, gastritis,
XX shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
XX

PS Claim 1; Page 46; 69pp; English.

XX
CC This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.
SQ Sequence 8 AA:

Query Match 84.48; Score 38; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. NO. 3.2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Caps 0;

OY 1 GRLCVDG 8
1 |||||
Db 1 grlcvdgd 8

RESULT 6
Y79114
ID Y79114 standard; Peptide; 8 AA.
XX
AC Y79114;

DT 05-JUN-2000 (first entry)

DE Peptide antagonist of zonulin.

XX
KW Zonulin; antagonist; zonula occludens toxin receptor;
KW blood-brain barrier; antiinflammatory; cerebroprotective;
KW neuroprotective; dermatological; antitumor; antiviral;
KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
KW gastrointestinal inflammation; therapy.

OS Synthetic.

PN WO200007609-A1.

PD 17-FEB-2000.

PF 28-JUL-1999; 99WO-US16663.

PR 03-AUG-1998; 98US-0127815.

PA (UYMA-) UNIV MARYLAND BALTIMORE.

PI Fasano A;

DR WPI; 2000-205565/18.

XX
PT New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
PS
PS Claim 1; Page 44; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
CC (2), one of 25 such peptides (see Y79105-29) of the invention,
CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
CC physiologically modulate the opening of mammalian tight junctions
CC (TJ). The peptide antagonists are based on a common motif of ZOT
CC and human zonulins, which is believed to be critical for receptor
CC binding. They can be prepared by chemical synthesis or by use of
CC recombinant DNA techniques. The peptide antagonists are used as an
CC antiinflammatory agents in the treatment of gastrointestinal
CC inflammation, where they bind to the ZOT receptor in the intestine
CC and yet does not physiologically modulate the opening of TJ in the
CC intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for
CC treating intestinal conditions that cause protein losing enteropathy
CC caused by infection, e.g. Clostridium difficile infection,
CC enterocolitis, shigellosis, viral gastroenteritis, parasite
CC infestation, bacterial overgrowth, whipple's disease, diseases with
CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
CC collagenous colitis, inflammatory bowel disease, diseases marked by
CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
CC correction of congenital heart disease with Fontan's operation,
CC mucosal diseases without ulceration, e.g. Menetrier's disease,
CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
CC e.g. systemic lupus erythematosus or food allergies, primarily to
CC milk.
SQ Sequence 8 AA:

Query Match 82.28; Score 37; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. NO. 3.2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Caps 0;

OY 1 GRLCVDG 8
1 |||||
Db 1 grlcvdgd 8

RESULT 7
R74094
ID R74094 standard; Protein; 742 AA.

AC R74094;

DT 04-JAN-1996 (first entry)

DE Human zona pellucida-2 (HZP-2) protein.

XX
KW Human zona pellucida; HZP-2; protein; vaccine; antigen;
KW contraceptive; ovary.

OS Homo sapiens.

PN JP07099974-A.

PD 18-APR-1995.

PF 05-OCT-1993; 93JP-0249404.

PR 05-OCT-1993; 93JP-0249404.

PA (TOFU) TONEN CORP.

PI WPI; 1995-182067/24.

DR N-PSDB; Q92254.

XX A DNA coding human zona pellucida-2 protein - used as a vaccine
 PT antigen
 XX
 PS Claim 1: Page 8-11; 11pp; Japanese.
 XX
 CC The human zona pellucida-2 (HZP-2) protein may be used as a vaccine
 CC antigen. It can be artificially synthesised using recombinant
 CC techniques. Partial sequences of the HZP-2 gene were subcloned, the
 CC sequences and primers used in cloning and subcloning are given in
 CC 092255-69.
 XX
 SO Sequence 742 AA;

Query Match 80.0%; Score 36; DB 16; Length 742;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRICVODG 8
 I I I I I I I
 Db 369 gelctcdg 376

RESULT 8
 R55206
 ID R55206 standard; Protein: 745 AA.
 XX
 AC R55206;
 XX
 DT 01-FEB-1995 (first entry)
 XX
 DE Human zona pellucida ZPA protein.
 XX
 KM Human; zona pellucida; ZPA; immunocontraception.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..745
 FT /label= human_ZPA
 XX
 PN W09411019-A.
 XX
 PD 26-MAY-1994.
 XX
 PF 06-NOV-1993; 93WO-US10851.
 XX
 PR 09-NOV-1992; 92US-0973341.
 PR 29-JAN-1993; 93US-0012990.
 XX
 PA (ZONA-) ZONAGEN INC.
 XX
 PI Harris JD, Hsu KT, Podolski JS;
 XX
 DR WPI: 1994-183156/22.
 DR N-PSDB; Q65616.
 XX
 PT Use of zona pellucida proteins and antibodies - for inducing
 PT reproducible transient infertility or permanent sterility in
 PT female mammals
 XX
 PS Claim 40: Page 126-128; 154pp; English.
 XX
 CC A commercially available human genomic DNA library consisting of
 CC 9-23kb inserts of DNA from placenta tissue of a male caucasian
 CC cloned into the lambda fix (RTM)II vector was screened with
 CC probes encoding porcine ZPA and ZPB proteins. Positive clones
 CC were analysed further by Southern hybridisation using the porcine
 CC probes and restriction analysis. Clones encoding human ZPA and ZPB
 CC proteins were identified (see Q65616 and Q65617, respectively).
 CC R55206 is the deduced amino acid sequence for human ZPA protein.
 XX

SO Sequence 745 AA;

Query Match 80.0%; Score 36; DB 15; Length 745;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRICVODG 8
 I I I I I I I
 Db 369 gelctcdg 376

RESULT 9
 Y42480
 ID Y42480 standard; Protein: 745 AA.
 XX
 AC Y42480;
 XX
 DT 15-FEB-2000 (first entry)
 XX
 DE Human zona pellucida ZPA protein.
 XX
 KM Contraceptive; porcine; pig; zona pellucida; mammal; lapine; canine;
 KM dog; rabbit; feline; cat; bovine; cow; human; cynomolgus monkey; ZPA;
 KM ZPB; ZPC; infertility; veterinary.
 XX
 OS Homo sapiens.
 XX
 PN US5981228-A.
 XX
 PD 09-NOV-1999.
 XX
 PF 07-JUN-1995; 95US-0484596.
 XX
 PR 09-NOV-1993; 93US-0149223.
 PR 09-NOV-1992; 92US-0973341.
 PR 29-JAN-1993; 93US-0012990.
 XX
 PA (ZONA-) ZONAGEN INC.
 XX
 PI Podolski JS, Hsu KT, Harris JD;
 XX
 DR WPI: 1999-633318/54.
 DR N-PSDB; Z22737.
 XX
 PT Mammalian zona pellucida proteins used to induce transient or permanent
 PT infertility -
 XX
 PS Example 11; Column 127-132; 84pp; English.
 XX
 CC This sequence represents the human zona pellucida ZPA protein. The
 CC invention relates to the isolation of novel nucleotide sequences encoding
 CC zona pellucida proteins from mammalian, especially porcine, lapine,
 CC canine, feline, bovine, human or cynomolgus monkey sources. The zona
 CC pellucida proteins (ZPA, ZPB, and ZPC) are used in the induction of
 CC transient or permanent infertility. At present the method is used in
 CC veterinary applications to induce transient or permanent infertility
 CC in porcine, lapine, canine, feline, bovine, and cynomolgus monkeys.
 XX
 SO Sequence 745 AA;

Query Match 80.0%; Score 36; DB 20; Length 745;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRICVODG 8
 I I I I I I I
 Db 369 gelctcdg 376

RESULT 10
 W81817

```

ID   W81817 standard; Protein; 745 AA.
XX
AC   W81817;
XX
DT   29-JAN-1999 (first entry)
XX
DE   Human ZPA protein.
XX
KW   ZPA; zona pellucida; infertility; sterility; immunononcontraceptive;
XX   vaccine; human.
XX
OS   Homo sapiens.
XX
PN   US5837497-A.
XX
PD   17-NOV-1998.
XX
PF   07-JUN-1995; 95US-0484993.
XX
PR   09-NOV-1993; 93US-0149223.
XX   09-NOV-1992; 92US-0973341.
XX   29-JAN-1993; 93US-0012990.
XX   07-JUN-1995; 95US-0484993.
XX
PA   (ZONA-) ZONAGEN INC.
XX
PI   Harris JD;
XX
DR   WPI; 1999-023447/02.
XX   N-PSDB; V64814.
XX
PT   Isolated zona pellucida DNA from different mammals - used to develop
XX   products which can be used for vaccination to induce transient
XX   infertility or permanent sterility in female mammals
XX
PS   Example 11; Column 133-138; 84pp; English.
XX
CC   This sequence represents a human ZPA protein isolated from zona
XX   pellucida. This protein can be used in a method for specifically
XX   inducing transient infertility or permanent sterility in a host
XX   animal by selective vaccination with specific zona pellucida proteins
XX   or immunononcontraceptively active fragments.
XX
SQ   Sequence 745 AA;

Query Match      80.0%; Score 36; DB 20; Length 745;
Best Local Similarity 75.0%; Pred. NO. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY   1 GRLCVQDG 8
    1 1 1 1 1
Db   369 gelctgdg 376

RESULT 11
ID   Y82215 standard; Protein; 745 AA.
XX
AC   Y82215;
XX
DT   12-JUN-2000 (first entry)
XX
DE   Human zona pellucida ZPA SEQ ID NO:43.
XX
KW   Zona pellucida; ZPA; ZPB; ZPC; infertility; permanent sterility;
XX   vaccine; immunononcontraceptive; contraceptive; immunostimulatory.
XX
OS   Homo sapiens.
XX
PN   US6027727-A.
XX
PD   22-FEB-2000.

```

```

XX
PF   09-NOV-1993; 93US-0149223.
XX
PR   09-NOV-1992; 92US-0973341.
XX   29-JAN-1993; 93US-0012990.
XX
PA   (ZONA-) ZONAGEN INC.
XX
PI   Podolski JS, Hsu KT, Harris JD;
XX
DR   WPI; 2000-269144/23.
XX   N-PSDB; 295674.
XX
PT   Inducing reproducible transient or permanent infertility in a mammal
XX   comprises administration of homologous and/or heterologous mammalian
XX   species zona pellucida proteins
XX
PS   Example 11; Column 135-138; 85pp; English.
XX
CC   The present invention describes a method for inducing reproducible
XX   transient infertility in a female mammal, including humans, by selective
XX   administration of homologous and/or heterologous mammalian species
XX   zona pellucida (ZP) proteins or immunononcontraceptively active fragments.
XX   Also described are: (1) a method for inducing transient infertility in a
XX   mammal by administering directly an antibody directed to a ZPB or an
XX   immunologically active fragment selected from feline, bovine,
XX   cynomolgous monkey or human ZPB polypeptides; (2) an isolated, purified
XX   recombinant feline, bovine, cynomolgous monkey or human ZPB polypeptide
XX   or immunologically active fragment; and (3) a fusion protein comprising
XX   a ZPB polypeptide which is conjugated with a compound selected from
XX   keyhole limpet haemocyanin, muramyl dipeptide, histidine-tag, beta-gal,
XX   or palmitic acid where the fusion protein remains effective to stimulate
XX   production of antibodies that recognise a ZPB polypeptide. The method is
XX   useful for providing transient or permanent infertility or sterility in
XX   humans and animals. The present sequence represents human ZPA, which is
XX   used in the exemplification of the present invention.
XX
SQ   Sequence 745 AA;

Query Match      80.0%; Score 36; DB 21; Length 745;
Best Local Similarity 75.0%; Pred. NO. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY   1 GRLCVQDG 8
    1 1 1 1 1
Db   369 gelctgdg 376

RESULT 12
ID   Y52689 standard; Protein; 745 AA.
XX
AC   Y52689;
XX
DT   14-MAR-2000 (first entry)
XX
DE   Human oocyte zona pellucida protein ZPA.
XX
KW   Zona pellucida; ZPA; ZPB; ZPC; purified; mammalian; glycoprotein;
XX   target; immunononcontraction; vaccine; antibody; transient; infertility;
XX   controllable; predictable; permanent; sterility; side effect.
XX
OS   Homo sapiens.
XX
PN   US6001599-A.
XX
PD   14-DEC-1999.
XX
PF   02-JUN-1995; 95US-0458731.
XX
PR   09-NOV-1993; 93US-0149223.
XX   09-NOV-1992; 92US-0973341.

```


PR 29-JAN-1993; 93US-0012990.
XX (ZONA-) ZONAGEN INC.
PA Podolski JS, Hsu KT, Harris JD;
XX WPI; 2000-061880/05.
DR
XX
PT Isolated DNA encoding mammalian zona pellucida proteins useful for
PT inducing transient sterility -
XX
PS Example 11; Columns 133-136; 84pp; English.
XX
CC This sequence represents a mammalian zona pellucida protein, ZPA.
CC The zona pellucida is a complex matrix surrounding the mammalian
CC oocyte, formed of glycoproteins secreted by ovarian cells. The
CC invention relates to zona pellucida proteins ZPA, ZPB and ZPC, which
CC are useful as targets for immunosuppressive vaccines. Polynucleotides
CC encoding these proteins can be used for recombinant protein production,
CC and as probes in hybridisation methods for the isolation of
CC polynucleotides encoding zona pellucida proteins from other mammalian
CC species. Administration of zona pellucida proteins to a host animal
CC results in a specific immunosuppressive effect. Administration of
CC purified ZPA and/or ZPB induces transient infertility in females.
CC Fertility can be maintained in an on or off condition in a controllable/
CC predictable fashion. Administration of ZPC induces permanent sterility.
CC Use of a purified zona pellucida protein rather than a complex mixture
CC of zona pellucida proteins reduces the potential for unwanted side-
CC effects which may be harmful.
XX
SQ Sequence 745 AA:
OY 1 GRLCVDG 8
DB 369 gelctcdg 376
Query Match 80.0%; Score 36; DB 21; Length 745;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
RESULT 13
Y52180
ID Y52180 standard; Protein; 745 AA.
XX
AC Y52180;
XX
DT 09-FEB-2000 (first entry)
XX
DE Human zona pellucida protein B (ZPB) amino acid sequence.
XX
KW Zona pellucida protein; ZPA; ZPB; ZPC; oocyte; ovarian cell; antibody;
KW fusion protein; vaccine; contraceptive; fertilization; fertility;
KW immunosuppressive.
XX
OS Homo sapiens.
XX
PN US5989550-A.
XX
PD 23-NOV-1999.
XX
PF 07-JUN-1995; 95US-0480150.
XX
PR 09-NOV-1993; 93US-0149223.
PR 09-NOV-1992; 92US-0973341.
PR 29-JAN-1993; 93US-0012990.
XX
PA (ZONA-) ZONAGEN INC.
XX
PI Podolski JS, Hsu KT, Harris JD;
XX
DR WPI; 2000-022762/02.

DR N-PSDB; 237821.
XX
PT Isolated and purified mammalian zona pellucida proteins useful in
PT methods of immunosuppression -
XX
PS Example 10; Column 133-136; 84pp; English.
XX
CC This is the human zona pellucida protein B (ZPB) amino acid sequence.
CC The invention relates to isolated and purified zona pellucida proteins
CC from mammals, specifically canine, feline and bovine ZPA or their
CC biologically active fragments also porcine and cynomolgus monkey ZPA. The
CC zona pellucida is a complex matrix surrounding the mammalian oocyte,
CC formed of glycoproteins secreted by ovarian cells. The previously
CC mentioned ZPA proteins can be used to produce a fusion protein to
CC stimulate production of antibodies in a mammal that recognize a ZPA
CC polypeptide of the mammal. The ZPA polypeptides are useful as vaccines
CC for selectively affecting transient infertility in mammals. ZP has an
CC important role in maintaining the oocyte and in oocyte-sperm interactions
CC and so is a target for a contraceptive agent which interferes with the
CC fertilization process. Providing a specific polypeptide reduces the need
CC for a complex mixture of many ZP proteins which may create unwanted and
CC harmful side effects. The duration of transient fertility is controllable
CC and can be maintained in a controllable and/or predictable fashion.
XX
SQ Sequence 745 AA:
OY 1 GRLCVDG 8
DB 369 gelctcdg 376
Query Match 80.0%; Score 36; DB 21; Length 745;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
RESULT 14
Y52984
ID Y52984 standard; Protein; 745 AA.
XX
AC Y52984;
XX
DT 14-FEB-2000 (first entry)
XX
DE Human zona pellucida ZPA protein sequence.
XX
KW Zona pellucida; ZPA; ZPB; ZPC; infertility; sterility; vaccine;
KW immunosuppressive; contraception; immunisation.
XX
OS Homo sapiens.
XX
PN US5976545-A.
XX
PD 02-NOV-1999.
XX
PF 07-JUN-1995; 95US-0484158.
XX
PR 09-NOV-1993; 93US-0149223.
PR 09-NOV-1992; 92US-0973341.
PR 29-JAN-1993; 93US-0012990.
XX
PA (ZONA-) ZONAGEN INC.
XX
PI Podolski JS, Hsu KT, Harris JD;
XX
DR WPI; 2000-037080/03.
DR N-PSDB; 233271.
XX
PT New recombinant mammalian zona pellucida C proteins, for induction of
PT female sterility -
XX
XX Example 11; Column 133-136; 87pp; English.

CC The present invention describes recombinant zona pellucida (ZP) proteins,
 CC specifically ZPC proteins. Also described are fusion proteins of ZPC with
 CC any of keyhole limpet haemocyanin, muramyl dipeptide, histidine tag,
 CC beta-galactosidase or palmitic acid, capable of stimulating mammalian
 CC antibodies that recognise mammalian ZPC. Recombinant ZPC proteins are
 CC used to immunise animals to induce transient infertility or permanent
 CC female sterility, in humans or other mammals. Z33243 to Z33254, Z33270 to
 CC Z33274 and Z33285 encode mammalian ZP proteins given in Y52971 to Y52988
 CC from the present invention. Z33255 to Z33269 and Z33275 to Z33284
 CC represent oligonucleotides used in the exemplification of the present
 CC invention.
 CC
 SQ Sequence 745 AA:
 CC
 Query Match 80.0%; Score 36; DB 21; Length 745;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GRICVDPG 8
 Db 369 gclctgdg 376
 CC
 RESULT 15
 Y79105
 ID Y79105 standard; Peptide; 8 AA.
 XX Y79105;
 AC
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Peptide antagonist of zonulin.
 XX
 KW Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antitumor; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.
 XX
 OS Synthetic.
 XX
 PN WO200007609-A1.
 XX
 PD 17-FEB-2000.
 XX
 PF 28-JUL-1999; 99WO-US16683.
 XX
 PR 03-AUG-1998; 98US-0127815.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Fasano A;
 XX
 DR WPI; 2000-205565/18.
 XX
 PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
 XX
 PS Claim 1; Page 41; 69pp; English.
 XX
 CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine

CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC coeliac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.
 CC
 SQ Sequence 8 AA:
 CC

Query Match 77.8%; Score 35; DB 21; Length 8;
 Best Local Similarity 75.0%; Pred. No. 3.2e+05;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRICVDPG 8
 Db 1 gricvdpq 8

Search completed: June 13, 2001, 14:14:33
 Job time: 374 sec

2001

This Page Blank (uspto)

**BEST COPY AVAILABLE
THIS PAGE IS BLANK (uspto)**

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:10:39 ; Search time 130.61 Seconds

(without alignments)
4.209 Million cell updates/sec

Title: PCT-US01-05825A-6

Perfect score: 45

Sequence: 1 GRICVODG 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	399	2	B82197 zona occludens tox
2	45	100.0	399	2	A43864 zonula occludens t
3	36	80.0	676	2	A45984 sperm-binding glyco
4	36	80.0	745	1	A48833 sperm-binding glyco
5	35	77.8	417	2	T05207 hypothetical prote
6	35	77.8	1677	2	T14267 Xln protein, stage
7	34	75.6	296	2	T26723 hypothetical prote
8	34	75.6	475	2	T01352 hypothetical prote
9	34	75.6	524	2	JN0606 ATP-stimulated glu
10	34	75.6	715	2	S70397 zona pellucida gly
11	34	75.6	716	2	S70398 zona pellucida gly
12	34	75.6	722	2	B71376 hypothetical prote
13	34	75.6	796	2	T23238 hypothetical prote
14	33	73.3	196	2	F71525 hypothetical prote
15	33	73.3	196	2	H81681 maf protein TC0628
16	33	73.3	293	2	B26637 neurogenic repetit
17	33	73.3	340	2	S69194 N4-Ibeta-N-acetyl
18	33	73.3	447	2	T34992 probable lipoprote
19	33	73.3	502	1	K1ECGL glycerol kinase (E
20	33	73.3	516	2	B81684 exodeoxyribonuclea
21	33	73.3	516	2	F71529 probable exodoxyl
22	33	73.3	967	2	T20336 hypothetical prote
23	33	73.3	971	2	T00394 GTP-binding protei
24	33	73.3	978	2	T00336 hypothetical prote
25	33	73.3	1012	2	T41940 DNA polymerase - h
26	33	73.3	2108	2	H70819 probable polyketid
27	33	73.3	2139	2	A35672 crumbs protein - f
28	32	71.1	122	2	C65009 hypothetical prote
29	32	71.1	322	2	S41400 aspartic proteins

30	32	71.1	328	2	I55450
31	32	71.1	328	2	UC4800
32	32	71.1	363	2	A83177
33	32	71.1	440	2	T43198
34	32	71.1	457	2	C75266
35	32	71.1	825	1	GLHO
36	32	71.1	857	2	T14471
37	32	71.1	1018	1	S73720
38	31	68.9	190	2	S39485
39	31	68.9	224	2	C72010
40	31	68.9	309	1	B64041
41	31	68.9	413	2	JE0142
42	31	68.9	433	2	A70536
43	31	68.9	443	2	T00419
44	31	68.9	445	2	S58282
45	31	68.9	445	2	T48072

ALIGNMENTS

RESULT 1
B82197
zona occludens toxin VC1458 [Imported] - Vibrio cholerae (group O1 strain N16961)
C:Species: Vibrio cholerae
C.Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C.Accession: B82197
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qian, H.; Dragol, I.; Sellers
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A.Reference number: AB2035; M0ID:20406833
A.Accession: B82197
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <HEI>
A:Cross-references: GB:AE004224; GB:AE003852; NID:99655952; PIDN:AAF94615.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1458
A:Map position: 1

Query Match 100.0%; Score 45; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GRICVODG 8
Db 291 GRICVODG 298

RESULT 2
A43864
zonula occludens toxin - Vibrio cholerae
C:Species: Vibrio cholerae
C.Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C.Accession: A43864
R:Baudry, B.; Pasano, A.; Kelsey, J.; Kaper, J.B.
Infect. Immun. 60, 428-434, 1992
A>Title: Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.
A:Reference number: A43864
A:Accession: A43864
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <BAU>
A:Cross-references: GB:M83563; NID:9155314; PIDN:AAA27582.1; PID:9155315
A:Note: sequence extracted from NCBI backbone (NCBIN:77488, NCBIIP:77491)

Query Match 100.0%; Score 45; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.29;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GRICVODG 8
|||||
Db 291 GRICVODG 298

RESULT 3
A:Accession: A45984
A:Species: sperm-binding glycoprotein ZP2 precursor - rabbit (fragment)
N:Alternate names: 75k zona pellucida glycoprotein; zona pellucida protein 2; zona pellucida
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 03-May-1994 #sequence_revision 05-Jan-1996 #text_change 05-Jan-1996
C:Accession: A45984
R:Lee, V.H.; Schwobel, E.; Prasad, S.; Cheung, P.; Timmons, T.M.; Cook, R.; Dunbar, B.S.
J. Biol. Chem. 268, 12412-12417, 1993
A:Title: Identification and structural characterization of the 75-kDa rabbit zona pellucida
A:Reference number: A45984; MUID:93286072
A:Accession: A45984
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-676 <LEP>
A:Cross-references: GB:112167
C:Superfamily: sperm-binding glycoprotein ZP2; ZP domain homology
C:Keywords: glycoprotein; oocyte; transmembrane protein
F:329-589/Domain: ZP domain homology <ZPH>

Query Match 80.0%; Score 36; DB 2; Length 676;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
|||||
Db 327 GELCTODG 334

RESULT 4
A:Accession: A48833
A:Species: sperm-binding glycoprotein ZP2 precursor - human
N:Alternate names: 75k zona pellucida glycoprotein; zona pellucida protein 2; zona pellucida
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Sep-1999
C:Accession: A48833
R:Liang, L.F.; Dean, J.
Dev. Biol. 156, 389-408, 1993
A:Title: Conservation of mammalian secondary sperm receptor genes enables the promoter
A:Reference number: A48833; MUID:93215931
A:Accession: A48833
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-745 <LTA>
A:Cross-references: GB:M90366; NID:g292939; PIDN:AAA61335.1; PID:g2929340
A:Experimental source: ovary
A:Note: sequence extracted from NCBI backbone (NCBIN:129165, NCBIIP:129166)
C:Genetics:
A:Gene: GDB:ZP2
A:Cross-references: GDB:6278878; OMIM:182888
C:Superfamily: sperm-binding glycoprotein ZP2; ZP domain homology
C:Keywords: glycoprotein; oocyte; transmembrane protein
F:371-631/Domain: ZP domain homology <ZPH>

Query Match 80.0%; Score 36; DB 1; Length 745;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
|||||
Db 369 GELCTODG 376

RESULT 5

T05207
hypothetical protein F24J7.30 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Oct-1999
C:Accession: T05207
R:Bevan, M.; Vitale, D.; Liguori, R.; Argiriou, A.; De Simone, V.; Bancroft, I.; Mewe
submitted to the Protein Sequence Database, February 1998
A:Reference number: Z13403
A:Accession: T05207
A:Molecule type: DNA
A:Residues: 1-417 <BEV>
A:Cross-references: EMBL:AL021768; GSPDB:GN00062; ATSP:F24J7.30
A:Experimental source: cultivar Columbia; BAC clone F24J7
C:Genetics:
A:Gene: ATSP:F24J7.30
A:Map position: 4
A:Introns: 33/1; 169/3

Query Match 77.8%; Score 35; DB 2; Length 417;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRICVOD 7
|||||
Db 373 GRICVOD 379

RESULT 6
T14267
Xin protein, stage early embryo - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14267
R:Wang, D.Z.; Lin, J.J.C.
submitted to the EMBL Data Library, March 1998
A:Description: Involvement of a novel gene, Xin, in cardiac looping.
A:Reference number: Z17948
A:Accession: T14267
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1677 <MAN>
A:Cross-references: EMBL:AF051945; NID:g2970645; PID:g2970646; PIDN:AAC06023.1
A:Experimental source: cardiac muscle; stage early embryo

Query Match 77.8%; Score 35; DB 2; Length 1677;
Best Local Similarity 75.0%; Pred. No. 97;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRICVODG 8
|||||
Db 886 GELCTODG 893

RESULT 7
T26723
hypothetical protein Y39A1A.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T26723
R:Wall, M.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z20257
A:Accession: T26723
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-296 <WIL>
A:Cross-references: EMBL:AL031633; PIDN:CAA21012.1; GSPDB:GN00021; CESP:Y39A1A.11
A:Experimental source: clone Y39A1A
C:Genetics:
A:Gene: CESP:Y39A1A.11
A:Map position: 3

A:introns: 45/3; 63/1; 114/3; 143/3; 225/3
C:Superfamily: ribitol dehydrogenase: short-chain alcohol dehydrogenase homology

Query Match 75.6%; Score 34; DB 2; Length 296;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
| | | | |
DB 16 GALCVHGD 23

RESULT 8

T01352
hypothetical protein F6N15.2 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 14-May-1999
C:Accession: T01352
R:Ryan, E.; Edwards, J.; Pape, K.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of A. thaliana F6N15.
A:Reference number: 214297
A:Accession: T01352
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-475 <RYA>
A:Cross-references: EMBL:AF069299; NID:g3193311; PID:g3193319
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:introns: 194/1
A:Note: F6N15.2

Query Match 75.6%; Score 34; DB 2; Length 475;
Best Local Similarity 71.4%; Pred. No. 47;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLCVODG 8
| | | | |
DB 208 RICEVDG 214

RESULT 9

JN0606
ATP-stimulated glucocorticoid-receptor translocation promoter protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 21-Jul-2000
C:Accession: JN0606
R:Okamoto, K.; Hirano, H.; Isohashi, F.
Biochem. Biophys. Res. Commun. 193, 848-854, 1993
A:Title: Molecular cloning of rat liver glucocorticoid-receptor translocation promoter.
A:Reference number: JN0606; MUID:93312338
A:Accession: JN0606
A:Molecule type: mRNA
A:Residues: 1-524 <OKA>
A:Cross-references: GB:D16102; NID:g414352; PID:BA03677.1; PID:g433211
A:Experimental source: liver
C:Comment: This protein is a histone binding protein which increases the glucocorticoid-
C:Superfamily: xylolokinease
C:Keywords: ATP

Query Match 75.6%; Score 34; DB 2; Length 524;
Best Local Similarity 62.5%; Pred. No. 52;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRICVODG 8
| | | | |
DB 266 GCMCFODG 273

RESULT 10

S70397
zona pellucida glycoprotein A - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 20-Aug-1999
C:Accession: S70397
R:Harlis, J.D.; Hibler, D.W.; Fontenot, G.K.; Hsu, K.T.; Yurewicz, E.C.; Sacco, A.G.
DNA Seq. 4, 361-393, 1994
A:Title: Cloning and characterization of zona pellucida genes and cDNAs from a variety
A:Reference number: S70396; MUID:95143578
A:Accession: S70397
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-715 <HAR>
A:Cross-references: EMBL:U05779; NID:g458274; PID:AAA74386.1; PID:g458275
C:Superfamily: sperm-binding glycoprotein ZP2; ZP domain homology
F:368-628/Domain: ZP domain homology <ZPH>

Query Match 75.6%; Score 34; DB 2; Length 715;
Best Local Similarity 75.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
| | | | |
DB 366 GDLCTODG 373

RESULT 11

S70398
zona pellucida glycoprotein A - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 20-Aug-1999
C:Accession: S70398
R:Harlis, J.D.; Hibler, D.W.; Fontenot, G.K.; Hsu, K.T.; Yurewicz, E.C.; Sacco, A.G.
DNA Seq. 4, 361-393, 1994
A:Title: Cloning and characterization of zona pellucida genes and cDNAs from a variety
A:Reference number: S70396; MUID:95143578
A:Accession: S70398
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-716 <HAR>
A:Cross-references: EMBL:U05776; NID:g458268; PID:AAA74388.1; PID:g458269
C:Superfamily: sperm-binding glycoprotein ZP2; ZP domain homology
F:370-630/Domain: ZP domain homology <ZPH>

Query Match 75.6%; Score 34; DB 2; Length 716;
Best Local Similarity 75.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
| | | | |
DB 368 GDLCTODG 375

RESULT 12

B71376
hypothetical protein TP0022 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 24-Nov-1999
C:Accession: B71376
R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Rson, J.; Khalek, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uitterback, T.; M they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770
A:Accession: B71376
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-722 <COL>
A:Cross-references: GB:AE001187; GB:AE000520; NID:g3322273; PID:AA05022.1; PID:g332

A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0022
 C:Superfamily: syphilis spirochete hypothetical protein TP0022

Query Match 75.6%; Score 34; DB 2; Length 722;
 Best Local Similarity 75.0%; Pred. No. 69;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRLCVODG 8
 Db 305 GRLCVODG 312

RESULT 13
 T23238
 hypothetical protein T05H10.7 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T23238; T24560
 R:Lighting, J.
 A:Submitted to the EMBL Data Library, January 1995
 A:Reference number: Z19713
 A:Accession: T23238
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-796 <WIL>
 A:Cross-references: EMBL:Z47811; PIDN:CAA87788.1; GSPDB:GN00020; CESP:T05H10.7
 A:Experimental source: clone K02C4
 R:Thomas, K.
 A:Submitted to the EMBL Data Library, January 1995
 A:Reference number: Z19907
 A:Accession: T24560
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-796 <WIL>
 A:Cross-references: EMBL:Z47812; PIDN:CAA87796.1; GSPDB:GN00020; CESP:T05H10.7
 A:Experimental source: clone T05H10
 C:Genetics:
 A:Gene: CESP:T05H10.7
 A:Map position: 2
 A:Introns: 77/1; 128/2; 163/1; 249/3; 559/3; 561/3; 648/3; 776/3

Query Match 75.6%; Score 34; DB 2; Length 796;
 Best Local Similarity 75.0%; Pred. No. 76;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRLCVODG 8
 Db 198 GRLCVODG 205

RESULT 14
 F71525
 hypothetical protein CT349 - *Chlamydia trachomatis* (serotype D, strain UM3/Cx)
 C:Species: *Chlamydia trachomatis*
 C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
 C:Accession: F71525
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*
 A:Reference number: A71570; MUID:99000809
 A:Accession: F71525
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-196 <ARN>
 A:Cross-references: GB:AE001308; GB:AE001273; MID:g3328766; PIDN:AA67944.1; PID:g332877
 C:Genetics:
 A:Experimental source: serotype D, strain UM-3/Cx
 A:Gene: CT349
 C:Superfamily: septum formation protein maf

Query Match 73.3%; Score 33; DB 2; Length 196;
 Best Local Similarity 75.0%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRLCVODG 8
 Db 150 GRLCVODG 157

RESULT 15
 H81681
 maf protein TC0628 [imported] - *Chlamydia muridarum* (strain Nigg)
 C:Species: *Chlamydia muridarum*, *Chlamydia trachomatis* MOpn
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
 C:Accession: H81681
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke, C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg, C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of *Chlamydia trachomatis* MOpn and *Chlamydia pneumoniae* AR39
 A:Reference number: AB1500; MUID:20150255
 A:Accession: H81681
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-196 <TET>
 A:Cross-references: GB:AE002331; GB:AE002160; MID:g7190662; PIDN:AAF39457.1; PID:g719
 A:Experimental source: strain Nigg (MOpn)
 C:Genetics:
 A:Gene: TC0628
 C:Superfamily: septum formation protein maf

Query Match 73.3%; Score 33; DB 2; Length 196;
 Best Local Similarity 75.0%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRLCVODG 8
 Db 150 GRLCVODG 157

Search completed: June 13, 2001, 14:10:40
 Job time: 141 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:21:41 ; Search time 74.44 Seconds
(without alignments)
3.681 Million cell updates/sec

Title: PCT-US01-05825A-6
Perfect score: 45
Sequence: 1 GRICVQDG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	399	1 ZOT_VIBCH	P38442 Vibrio chol
2	37	82.2	1120	1 DPOL_RCMVM	O85428 rat cytoleg
3	36	80.0	666	1 ZP2_RABIT	P48829 oryctolagus
4	36	80.0	745	1 ZP2_HUMAN	O05996 homo sapien
5	36	80.0	745	1 ZP2_MACRA	O77726 macaca radi
6	34	75.6	524	1 GLPK_MOUSE	O64316 mus musculu
7	34	75.6	524	1 GLPK_MOUSE	O63060 rattus norv
8	34	75.6	715	1 ZP2_CANFA	P47383 canis fami
9	34	75.6	715	1 ZP2_FELCA	P47383 felis silve
10	34	75.6	722	1 Y022_TREPA	O83066 treponema p
11	34	75.6	775	1 SM3E_HUMAN	O15041 homo sapien
12	34	75.6	775	1 SM3E_MOUSE	P70275 mus musculu
13	34	75.6	796	1 YRST_CAEEL	O10003 caenorhabdi
14	34	75.6	1171	1 DPOL_HSVT1	O9yus3 herpesvirus
15	34	75.6	1171	1 DPOL_HSVT2	O9yus2 herpesvirus
16	33	73.3	327	1 O6A1_HUMAN	O95222 homo sapien
17	33	73.3	340	1 ASBP_FLAME	O47898 flavobacter
18	33	73.3	501	1 GLPK_ECOLI	P08859 escherichia
19	33	73.3	1012	1 DPOL_HSVT1	P52342 human herpe
20	33	73.3	2139	1 CRB_DROME	P10040 drosophila
21	32	71.1	122	1 YFDO_ECOLI	P76511 escherichia
22	32	71.1	328	1 P2Y6_HUMAN	O15077 homo sapien
23	32	71.1	328	1 P2Y6_HUMAN	O63371 rattus norv
24	32	71.1	511	1 C4A2_DROME	O9yus8 drosophila
25	32	71.1	825	1 BGUS_HANAN	P06835 haemaphysa
26	32	71.1	1018	1 HMM1_MYCPN	O50365 mycoplasma
27	31	68.9	70	1 NUSG_CGDAB	P36262 citrus gree
28	31	68.9	309	1 YBKG_HAETN	P44298 haemophilus
29	31	68.9	480	1 GLTD_AOBR	O05756 azospirillum
30	31	68.9	496	1 PACR_MOUSE	P70205 mus musculu
31	31	68.9	523	1 PACR_MOUSE	P32215 rattus norv
32	31	68.9	785	1 SM3E_CHICK	O42237 gallus gall
33	31	68.9	876	1 BGIL_SACFI	P22506 saccharomyc

34	31	68.9	880	1 BGL2_SACFI	P22507 saccharomyc
35	31	68.9	968	1 PKD2_HUMAN	O13563 homo sapien
36	31	68.9	1035	1 DPOL_RHOM6	O71121 rhesus cyto
37	31	68.9	1242	1 DPOL_HCMVA	P08546 human cytom
38	30	66.7	342	1 CYSF_SCHJA	P43157 schistosoma
39	30	66.7	393	1 LA2M_MYCSM	P21795 mycobacteri
40	30	66.7	406	1 LMP1_MOUSE	P11438 mus musculu
41	30	66.7	407	1 LMP1_MOUSE	P11438 mus musculu
42	30	66.7	421	1 DCDA_ZYMO	O92651 zymomonas m
43	30	66.7	446	1 CRB3_SCHPO	O10272 schizosacch
44	30	66.7	471	1 GLTD_ECOLI	P09832 escherichia
45	30	66.7	508	1 TRPE_SYNY3	P20170 synchocyst

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	399 AA.
ID	ZOT_VIBCH			
AC	P38442; Q9L706; Q9R3V6;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-2000 (Rel. 40, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	ZONA OCCLUDENS TOXIN (ZONULAR OCCLUDENS TOXIN).			
GN	ZOT OR VC1458.			
OS	Vibrio cholerae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.			
OX	NCBI_TaxID=666;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CLASSICAL INABA 569B;			
RX	MEDLINE=92112300; PubMed=1730472;			
RA	Baudry B., Fasano A., Kellely J., Kaper J.B.;			
RT	"Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.";			
RL	Infect. Immun. 60:428-434(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=KN1H002;			
RA	Shin H.J., Park Y.C., Kim Y.C.;			
RT	"Cloning and nucleotide sequence analysis of the virulence gene cassette from Vibrio cholerae KN1H002 isolated in Korea.";			
RL	Misamunthrag Holji 35:205-210(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=O139-TOR OGAMA;			
RA	Zhi-Yong H., Wei-Jie Z., Xiang-Fu W.;			
RT	"Cloning and Expression of zot Gene from Vibrio cholerae.";			
RL	Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=EL TOR 86015 / SEROTYPE O1;			
RA	Ken B., Liu Y.O., Qi G.M., Gao S.Y.;			
RT	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=EL TOR N16961 / SEROTYPE O1;			
RX	MEDLINE=20406833; PubMed=10952301;			
RA	Hedden R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,			
RT	Dooson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,			
RA	Gill S.R., Nelson K.E., Read T.D., Tetteilin H., Richardson D.,			
RA	Emmolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,			
RA	McDonald L., Ulfertback T., Fleishmann R.D., Nierman W.C., White O.,			
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,			
RA	Fraser C.M.;			
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.";			
RL	Nature 406:477-483(2000).			
RN	[6]			
RP	CHARACTERIZATION.			
RX	MEDLINE=91271365; PubMed=2052603;			
RA	Fasano A., Baudry B., Pumphlin D.W., Wasserman S.S., Tall B.D.,			

RA Kettley J.M., Kaper J.B.;
 RT Vibrio cholerae produces a second enterotoxin, which affects
 RL intestinal tight junctions.*;
 CC Proc. Natl. Acad. Sci. U.S.A. 88:5242-5246(1991).
 CC -1- FUNCTION: INCREASES THE PERMEABILITY OF THE SMALL INTESTINE MUCOSA
 CC BY AFFECTING THE STRUCTURE OF INTERCELLULAR TIGHT JUNCTIONS
 CC (ZONULA OCCLUDENS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M83563; AAA27582.1; -
 DR EMBL: AF175708; AAD51358.1; -
 DR EMBL: AF123049; AAD26854.1; -
 DR EMBL: AF220606; AAF29547.1; -
 DR EMBL: AE004224; AAF94615.1; -
 DR PIR: A43864; A43864.
 DR TIGR: VC1458; -
 DR Enterotoxin; Toxin.
 KW VARIANT 45 45 M -> I (IN STRAIN 569B).
 FT VARIANT 100 100 V -> A (IN STRAINS 569B AND 86015).
 FT VARIANT 272 272 V -> A (IN STRAIN 569B).
 FT VARIANT 281 281 V -> A (IN STRAIN 569B).
 FT VARIANT 349 349 V -> S (IN STRAIN 86015).
 FT VARIANT 381 381 K -> R (IN STRAIN 86015).
 FT CONFLICT 386 399 IKTENDRKGLNSIF -> VKREKESEIISPL (IN REF.
 FT SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;
 SO

Query Match 100.0%; Score 45; DB 1; Length 399;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GRLCVDG 8
 Db 291 GRLCVDG 298

RESULT 2
 DPOL_RCMVM STANDARD; PRT; 1120 AA.
 AC Q85428; 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA POLYMERASE (EC 2.7.7.7).
 GN UL54.
 OS Rat cytomegalovirus (strain Mastricht).
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Betaherpesvirinae; Muromegalovirus.
 CC NCBI_Taxid=79700;
 CC [1]
 CC SEQUENCE FROM N.A.
 RX MEDLINE=96335691; PubMed=8757999;
 RA Beuken E., Stobbe R., Bruggeman C.A., Vink C.;
 RT "Cloning and sequence analysis of the genes encoding DNA polymerase,
 RT glycoprotein B, ICP18.5 and major DNA-binding protein of rat
 RT cytomegalovirus.";
 RL J. Gen. Virol. 77:1559-1562(1996).
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
 CC N PYROPHOSPHATE + DNA(N).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U50550; AAC56433.1; -
 DR InterPro: IPR002064; -
 DR Pfam: PF00136; DNA_POL_B; 1.
 DR PRINTS: PR00106; DNAPOLB.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
 KW Transferase; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Nuclear protein.
 SO SEQUENCE 1120 AA; 124853 MW; 12FFAAE95EF54FC4 CRC64;
 S

Query Match 82.2%; Score 37; DB 1; Length 1120;
 Best Local Similarity 75.0%; Pred. No. 12;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GRLCVDG 8
 Db 493 GRLCVDG 500

RESULT 3
 ZP2_RABIT STANDARD; PRT; 666 AA.
 ID ZP2_RABIT
 AC P48829;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 2 (ZONA PELLUCIDA GLYCOPROTEIN
 DE (FRAGMENT)).
 DE ZP2 (ZONA PELLUCIDA PROTEIN A) (75 KDA ZONA PELLUCIDA PROTEIN)
 DE (FRAGMENT).
 GN ZP2 OR ZPA.
 GN Oryctolagus cuniculus (Rabbit).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 CC NCBI_Taxid=9986;
 CC [1]
 CC SEQUENCE FROM N.A.
 RX STRAIN=NEW ZEALAND WHITE; TISSUE=OVARY;
 RX MEDLINE=93286072; PubMed=7685342;
 RA Lee V.H., Schoebel E.D., Prasad S.V., Cheung P., Timmons T.M.,
 RA Cook R.G., Dunbar B.S.;
 RT "Identification and structural characterization of the 75-kDa rabbit
 RT zona pellucida protein.";
 RL J. Biol. Chem. 268:12412-12417(1993).
 CC -1- FUNCTION: ZP2 FORMS WITH ZP1 AND ZP3 THE ZONA PELLUCIDA, IN
 CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC ZP2 ACTS AS A SECONDARY SPERM RECEPTOR (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
 CC MATRIX.
 CC -1- PTM: PROTEOLYTICALLY CLEAVED AFTER FERTILIZATION, AND THIS
 CC MODIFICATION, ALONG WITH PRESUMED CHANGES IN ZP3 MAY PLAY AN
 CC IMPORTANT ROLE IN THE POSTFERTILIZATION BLOCK TO POLYSPERM.
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L12167; AAA31502.1; -
 DR InterPro: IPR001507; -
 DR Pfam: PF00100; zona_pellucida; 1.
 DR PROSITE: PS00682; ZP_DOMAIN; 1.
 KW Glycoprotein; Sulfatation; Sperm; Receptor; Transmembrane;

KW Extracellular matrix.
 FT NON_TER 1 1
 FT CHAIN 1 666
 FT DOMAIN <1 636
 FT TRANSMEM 637 656
 FT DOMAIN 657 666
 FT DOMAIN 318 585
 FT CARBOHYD 38 38
 FT CARBOHYD 73 73
 FT CARBOHYD 126 126
 FT CARBOHYD 171 171
 FT CARBOHYD 217 217
 FT CARBOHYD 241 241
 FT CARBOHYD 613 613
 SQ SEQUENCE 666 AA; 73644 MW; D6C8E2BA2D21020B CRC64;

Query Match
 Best Local Similarity 80.0%; Score 36; DB 1; Length 666;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRICVODG 8
 DB 317 GELCTQDG 324

RESULT 4
 ZP2_HUMAN
 ID ZP2_HUMAN STANDARD; PRT; 745 AA.
 AC Q05396;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 2 PRECURSOR (ZONA PELLUCIDA GLYCOPROTEIN ZP2) (ZONA PELLUCIDA PROTEIN A).
 GN ZP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=93215931; PubMed=8385033;
 RA Liang L.-F., Dean J.;
 RT "Conservation of mammalian secondary sperm receptor genes enables the promoter of the human gene to function in mouse oocytes.";
 RL Dev. Biol. 156:399-408(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99425270; PubMed=10493829;
 RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fushan J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslantes Maye A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
 RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";
 RL Genomics 60:295-308(1999).
 CC -1- FUNCTION: ZP2 FORMS WITH ZP1 AND ZP3 THE ZONA PELLUCIDA, IN WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR MATRIX.
 CC -1- TISSUE SPECIFICITY: OOCYTES.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING THE 2-WEEK GROWTH PHASE OF OOGENESIS, PRIOR TO OVULATION.
 CC -1- PTM: IS PROTEOLYTICALLY CLEAVED AFTER FERTILIZATION, AND THIS MODIFICATION, ALONG WITH PRESUMED CHANGES IN ZP3 MAY PLAY AN IMPORTANT ROLE IN THE POSTFERTILIZATION BLOCK TO POLYSPERMY.
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M90366; AAA6135.1; -
 DR EMBL: AF001550; AAB67599.1; -
 DR PIR: A48833; A48833.
 DR MIM: 182888; -
 DR InterPro: IP001507; -
 DR Pfam: PF00100; zona_pellucida; 1.
 DR PRINTS: PR00023; ZP_DOMAIN.1.
 DR PROSITE: PS00682; ZP_DOMAIN.1.
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane; Extracellular matrix.
 FT SIGNAL 1 38
 FT CHAIN 39 745
 FT DOMAIN 39 716
 FT TRANSMEM 717 736
 FT DOMAIN 737 745
 FT DOMAIN 370 637
 FT CARBOHYD 87 87
 FT CARBOHYD 105 105
 FT CARBOHYD 122 122
 FT CARBOHYD 223 223
 FT CARBOHYD 269 269
 FT CARBOHYD 400 400
 SQ SEQUENCE 745 AA; 82356 MW; 261AD79738F2CDD CRC64;

Query Match
 Best Local Similarity 80.0%; Score 36; DB 1; Length 745;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRICVODG 8
 DB 369 GELCTQDG 376

RESULT 5
 ZP2_MACRA
 ID ZP2_MACRA STANDARD; PRT; 745 AA.
 AC 077726;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 2 PRECURSOR (ZONA PELLUCIDA GLYCOPROTEIN ZP2) (ZONA PELLUCIDA PROTEIN A).
 GN ZP2.
 OS Macaca radiata (Bonnate monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OX NCBI_TaxID=9548;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=98250422; PubMed=9590540;
 RA Jethanandani P., Santhanam R., Gupta S.K.;
 RT "Molecular cloning and expression in Escherichia coli of cDNA encoding bonnet monkey (Macaca radiata) zona pellucida glycoprotein-zp2.";
 RL Mol. Reprod. Dev. 50:229-239(1998).
 CC -1- FUNCTION: ZP2 FORMS WITH ZP1 AND ZP3 THE ZONA PELLUCIDA, IN WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR MATRIX.
 CC -1- PTM: IS PROTEOLYTICALLY CLEAVED AFTER FERTILIZATION, AND THIS MODIFICATION, ALONG WITH PRESUMED CHANGES IN ZP3 MAY PLAY AN

```

CC CC IMPORTANT ROLE IN THE POSTFERTILIZATION BLOCK TO POLYSPERMY.
CC CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES
CC CC (BY SIMILARITY).
CC CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@sib-sib.ch).
-----
CC CC EMBL: Y10690; CAA71693.1; -
CC CC InterPro: IPR001507; -
CC CC Pfam: PF00100; zona_pellucida; 1.
CC CC PRINTS: PR00023; ZPELLUCIDA.
CC CC PROSITE: PS00682; ZP_DOMAIN; 1.
CC CC Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
CC CC Extracellular matrix.
CC CC SIGNAL 1 38
CC CC CHAIN 39 745 ZONA PELLUCIDA SPERM-BINDING PROTEIN 2.
CC CC DOMAIN 39 716 EXTRACELLULAR (POTENTIAL).
CC CC TRANSMEM 717 736 POTENTIAL.
CC CC DOMAIN 737 745 CYTOPLASMIC (POTENTIAL).
CC CC DOMAIN 370 637 ZP.
CC CC CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC CARBOHYD 745 AA; 82710 MW; 731D9AFAD3EE028 CRC64;
CC CC SEQUENCE
-----
Query Match. 80.0%; Score 36; DB 1; Length 745;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
-----
CC CC 1 GRLCVQDG 8
CC CC 1 1 1 1 1 1
CC CC Db 369 GELCTQDG 376
-----
RESULT 6
GLPK_MOUSE STANDARD; PRT; 524 AA.
AC 064516;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE)
DE (GLYCEROKINASE) (GK).
GN GK OR GYK.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=97038697; PubMed=8884278;
RA Hug A.H., Lovell R.S., Sampson W.J., Decker W.K., Dinulos M.B.,
RA Distche C.M., Craigen W.J.;
RT "Isolation, mapping, and functional expression of the mouse X
RT chromosome glycerol kinase gene.";
RL Genomics 36:530-534(1996).
CC CC -1- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND
CC CC METABOLISM (BY SIMILARITY).
CC CC -1- CATALYTIC ACTIVITY: ATP + GLYCEROL = ADP + GLYCEROL 3-PHOSPHATE.
CC CC -1- PATHWAY: RATE-LIMITING STEP IN GLYCEROL UTILIZATION.
CC CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL SURFACE OR

```

```

CC CC CYTOPLASMIC (BY SIMILARITY).
CC CC -1- SIMILARITY: BELONGS TO THE FUCCOKINASE / GLUCONOKINASE /
CC CC GLYCEROKINASE / XYLUOKINASE FAMILY.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@sib-sib.ch).
-----
CC CC EMBL: U48403; AAC52824.1; -
CC CC HSSP: P08859; 1GLB.
CC CC MGD: MGI:106594; GYK.
CC CC InterPro: IPR000577; -
CC CC Pfam: PF00370; FGGY; 1.
CC CC PROSITE: PS00445; FGGY_KINASES_2; 1.
CC CC PROSITE: PS00933; FGGY_KINASES_1; 1.
CC CC Glycerol metabolism; Transferase; Kinase; ATP-binding.
CC CC NP_BIND 167 179 ATP (PROBABLY).
CC CC SEQUENCE 524 AA; 57458 MW; C3E1C56887683B55 CRC64;
-----
Query Match. 75.6%; Score 34; DB 1; Length 524;
Best Local Similarity 62.5%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
-----
CC CC 1 GRLCVQDG 8
CC CC 1 1 1 1 1 1
CC CC Db 266 GQWCFQDG 273
-----
RESULT 7
GLPK_RAT STANDARD; PRT; 524 AA.
AC Q63060;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL 3-PHOSPHOTRANSFERASE)
DE (GLYCEROKINASE) (GK) (ATP-STIMULATED GLUCOCORTICOID-RECEPTOR
DE TRANSLLOCATION PROMOTER) (ASTP).
GN GK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Liver;
RX MEDLINE=93312338; PubMed=8323560;
RA Okamoto K., Hirano H., Isohashi F.;
RT "Molecular cloning of rat liver glucocorticoid-receptor translocation
RT promoter.";
RL Biochem. Biophys. Res. Commun. 193:848-854(1993).
CC CC -1- FUNCTION: INCREASES THE BINDING OF ACTIVATED GLUCOCORTICOID-
CC CC RECEPTOR TO NUCLEI IN THE PRESENCE OF ATP.
CC CC -1- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND
CC CC METABOLISM (BY SIMILARITY).
CC CC -1- CATALYTIC ACTIVITY: ATP + GLYCEROL = ADP + GLYCEROL 3-PHOSPHATE.
CC CC -1- PATHWAY: RATE-LIMITING STEP IN GLYCEROL UTILIZATION.
CC CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL SURFACE OR
CC CC CYTOPLASMIC (BY SIMILARITY).
CC CC -1- SIMILARITY: BELONGS TO THE FUCCOKINASE / GLUCONOKINASE /
CC CC GLYCEROKINASE / XYLUOKINASE FAMILY.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

CC or send an email to license@lsb-sib.ch.
 CC EMBL: D16102; BAA03677.1; -.
 DR HSSP: P08859; 1GLB.
 DR InterPro: IPR000577; -.
 DR Pfam: PF00370; FGGV; 1.
 DR PROSITE: PS00933; FGGV_KINASES_1; 1.
 DR PROSITE: PS00445; FGGV_KINASES_2; 1.
 KM GlyceroL metabolism; Transferase; Kinase; ATP-binding.
 FT N-BIND 167 179 ATP (PROBABLE).
 SQ SEQUENCE 524 AA: 57477 MW: A620C296D32B6B84 CRC64;

Query Match
 Best Local Similarity 75.6%; Score 34; DB 1; Length 524;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRLCVODG 8
 Db 266 GCMCFODG 273

RESULT 8
 ID ZP2_CANFA STANDARD; PRT; 715 AA.
 AC P47983;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE ZONA PELUCIDA SPERM-BINDING PROTEIN 2 PRECURSOR (ZONA PELUCIDA
 DE GLYCOPROTEIN ZP2) (ZONA PELUCIDA PROTEIN A).
 GN ZP2 OR ZPA.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE-95143578; PubMed-7841460;
 RA Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
 RA Sacco A.G.;
 RT "Cloning and characterization of zona pellucida genes and cDNAs from
 RT a variety of mammalian species: the ZPA, ZPB and ZPC gene families";
 RL DNA Seq. 4:361-393(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RC Okazaki Y., Isojima S., Sugimoto M.;
 RA Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC
 CC EMBL: U05779; BAA74386.1; -.
 DR EMBL: D45069; BAA08097.1; -.
 DR InterPro: IPR001507; -.
 DR Pfam: PF00100; zona_pellucida; 1.
 DR PRINTS: PR00023; ZPELUCIDA.
 DR PROSITE: PS00682; ZP_DOMAIN; 1.
 KM Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
 KM Extracellular matrix.
 FT SIGNAL 1 38
 FT CHAIN 39 715
 FT DOMAIN 39 684
 FT TRANSMEM 685 705
 FT DOMAIN 706 715
 FT DOMAIN 367 634
 FT BY SIMILARITY
 FT ZONA PELUCIDA SPERM-BINDING PROTEIN 2.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT ZP.

FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 220 220 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 266 266 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 15 15 R -> W (IN REF. 2).
 FT CONFLICT 292 292 R -> A (IN REF. 2).
 FT CONFLICT 328 328 L -> P (IN REF. 2).
 FT CONFLICT 599 599 S -> A (IN REF. 2).
 SQ SEQUENCE 715 AA: 79938 MW: 508D6DE804F4DC5C CRC64;

Query Match
 Best Local Similarity 75.0%; Score 34; DB 1; Length 715;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GRLCVODG 8
 Db 366 GDLCTODG 373

RESULT 9
 ID ZP2_FELCA STANDARD; PRT; 716 AA.
 AC P47984;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ZONA PELUCIDA SPERM-BINDING PROTEIN 2 PRECURSOR (ZONA PELUCIDA
 DE GLYCOPROTEIN ZP2) (ZONA PELUCIDA PROTEIN A).
 GN ZP2 OR ZPA.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE-95143578; PubMed-7841460;
 RA Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
 RA Sacco A.G.;
 RT "Cloning and characterization of zona pellucida genes and cDNAs from
 RT a variety of mammalian species: the ZPA, ZPB and ZPC gene families";
 RL DNA Seq. 4:361-393(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RC Okazaki Y., Isojima S., Sugimoto M.;
 RA Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ZP2 FORMS WITH ZP1 AND ZP3 THE ZONA PELUCIDA, IN
 CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
 CC MATRIX.
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC
 CC EMBL: U05776; BAA74388.1; -.
 DR EMBL: D45067; BAA08095.1; -.
 DR InterPro: IPR001507; -.
 DR Pfam: PF00100; zona_pellucida; 1.
 DR PRINTS: PR00023; ZPELUCIDA.
 DR PROSITE: PS00682; ZP_DOMAIN; 1.
 KM Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;

```
KW Extracellular matrix.
FT SIGNAL 1 38
FT CHAIN 39 716
FT DOMAIN 39 686
FT TRANSMEM 687 707
FT DOMAIN 708 716
FT DOMAIN 369 636
FT CARBOHYD 87 87
FT CARBOHYD 96 96
FT CARBOHYD 222 222
FT CARBOHYD 268 268
FT CARBOHYD 351 351
FT CARBOHYD 397 397
FT CONFLICT 483 483
FT CONFLICT 637 637
FT CONFLICT 637 637
SQ SEQUENCE 716 AA; 80135 MW; C5745496E82CB671 CRC64;

Query Match
Best Local Similarity 75.6%; Score 34; DB 1; Length 716;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRICVODG 8
Db 368 GDLCTODG 375

RESULT 10
Y022_TREPA STANDARD; PRT; 722 AA.
AC 083066;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE HYPOTHEICAL PROTEIN TP0022.
GN TP0022.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN NCBTAXID=160;
RP SEQUENCE FROM N.A.
RC MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardson J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uitterback T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Harch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RT Science 281:375-388(1998).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001187; AAC65022.1; -
DR TIGR; TP0022; -
KW Hypothetical protein.
SQ SEQUENCE 722 AA; 80089 MW; BBA7E917E310F560 CRC64;
```

```
Db 305 GLICVODG 312
OY 1 GRICVODG 7
Db 267 GRICVODG 273

RESULT 11
SM3E_HUMAN STANDARD; PRT; 775 AA.
ID SM3E_HUMAN
AC 015041;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE SEMAPHORIN 3E PRECURSOR.
GN SEMAPHORIN 3E OR KIA00331.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN NCBTAXID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro."
RL DNA Res. 4:141-150(1997).
RN [2]
RP SEQUENCE OF 1-38 FROM N.A.
RA Holmes A., Johnson D., Elliot G., Burkhardt J.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB002329; BAA20789.1; -
DR EMBL; AC004954; AAC69513.1; -
DR InterPro: IPR001627; -
DR InterPro: IPR003066; -
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF00047; Ig; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 775
FT CHAIN 1 775
FT DOMAIN 241 540
FT DOMAIN 647 736
FT DOMAIN 737 770
FT DISULFID 654 729
FT CARBOHYD 44 44
FT CARBOHYD 126 126
FT CARBOHYD 330 330
FT CARBOHYD 595 595
FT CARBOHYD 596 596
SQ SEQUENCE 775 AA; 89227 MW; CD6079C1AE48F779 CRC64;
```

```
Query Match
Best Local Similarity 75.6%; Score 34; DB 1; Length 775;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRICVODG 7
Db 267 GRICVODG 273
```



```

RESULT 12
SM3E_MOUSE STANDARD: PRT: 775 AA.
AC P70275: 009078; 009079;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE SEMAPHORIN 3E PRECURSOR (SEMAPHORIN H) (SEMA H).
GN SEMA3E OR SEMAH OR SEMH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=98175564; PubMed=9515811;
RA Christensen C.R.L., Klingelhofer J., Tarabkina S., Hultgaard E.F.,
RA Kramerov D., Lukandlin E.;
RT "Transcription of a novel mouse semaphorin gene, M-semah, correlates
RT with the metastatic ability of mouse tumor cell lines.";
RL Cancer Res. 58:1238-1244(1998).
RN [2]
RP REVISIONS.
RA Christensen C.R.L.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED IN EMBRYONIC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: DEVELOPING LUNGS, DEVELOPING SKELETAL
CC ELEMENTS, AND VENTRAL HORNS OF THE DEVELOPING NEURAL TUBE.
CC CORRELATES POSITIVELY WITH TUMOR PROGRESSION.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z80941; CAB02590.1; -
DR EMBL: Z93947; CAB07987.1; ALT_SEQ.
DR EMBL: Z93948; CAB07988.1; ALT_SEQ.
DR MGD: MGI:1340034; Sem3e.
DR InterPro: IPR001627; -
DR InterPro: IPR003006; -
DR Pfam: PF01403; Sema; 1.
DR Pfam: PF00047; Ig; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 775
FT CHAIN ? SEMAPHORIN 3E.
FT DOMAIN 241 540
FT DOMAIN 647 736
FT DOMAIN 737 770
FT DISULFID 654 729
FT CARBOHYD 44 44
FT CARBOHYD 126 126
FT CARBOHYD 175 175
FT CARBOHYD 330 330
FT CARBOHYD 596 596
FT SEQUENCE 775 AA; 89503 MW; BA41690093D28F40 CRC64;

```

```

DB 267 GRLCYND 273
RESULT 13
YRS7_CAEEI STANDARD: PRT: 796 AA.
AC Q10003; 1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 90.8 KDA PROTEIN T05H10.7 IN CHROMOSOME II.
GN T05H10.7
GN Caenorhabditis elegans.
OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Lightning J., Thomas K.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z47811; CA887788.1; -
DR EMBL: Z47812; CA887788.1; JOINED.
DR EMBL: Z47812; CA887796.1; -
DR EMBL: Z47811; CA887796.1; JOINED.
DR WormPep: T05H10.7; CE03637.
KW Hypothetical protein.
FT DOMAIN 38 42
FT DOMAIN 67 70
FT DOMAIN 524 527
FT SEQUENCE 796 AA; 90831 MW; 7BDFF8E0A4D2AA9F1 CRC64;

```

Query Match Best Local Similarity 75.6% Score 34; DB 1; Length 796;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 GRLCYNDG 8
DB 198 GRLCYNDG 205

```

```

RESULT 14
DPOL_HSVT1 STANDARD: PRT: 1171 AA.
AC Q9YUS3;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DNA POLYMERASE (EC 2.7.7.7).
GN DPOL.
OS Herpesvirus tupaia (Strain 1) (THV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae;
OX NCBI_TaxID=132677;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99094630; PubMed=9880021;
RA Springfield C., Tidona C.A., Kehm R., Bahr U., Darai G.;
RT "Identification and characterization of the Tupaia herpesvirus DNA
RL polymerase gene.";
RL J. Gen. Virol. 79:3049-3053(1998).
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC N PYROPHOSPHATE + DNA(N).

```

```

CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF074327; AAD08666.1; -.
CC InterPro: IPR002064; -.
CC Pfam: PF00136; DNA_POL_B: 2.
CC PRINTS: PR00106; DNAPOLB.
CC PROSITE: PS00116; DNA_POLYMERASE_B: 1.
CC Transferrase: DNA-directed DNA polymerase; DNA replication;
CC DNA-binding; Nuclear protein.
CC SEQUENCE 1171 AA; 128622 MW; CDC0480FEACFC7EC CRC64;
SQ

```

```

Query Match          75.6%; Score 34; DB 1; Length 1171;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 GRLCYOD 7
   111111
Db 515 GRLCYOD 521

```

```

RESULT 15
DPOL_HSVT2          STANDARD; PRT; 1171 AA.
AC 09YUS2:
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DNA POLYMERASE (EC 2.7.7.7).
GN DPOL.
OS Herpesvirus tupaia (Strain 2) (THV-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=132678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99094630; PubMed=9880021;
RA Springfield C., Tidona C.A., Kehm R., Bahr U., Darai G.;
RT "Identification and characterization of the Tupaia herpesvirus DNA
RT polymerase gene.";
RL J. Gen. Virol. 79:3049-3053(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99319892; PubMed=10392721;
RA Bahr U., Springfield C., Tidona C.A., Darai G.;
RT "Structural organization of a conserved gene cluster of Tupaia
RT herpesvirus encoding the DNA polymerase, glycoprotein B, a probable
RT processing and transport protein, and the major DNA binding protein.";
RL Virus Res. 60:123-136(1999).
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC N PYROPHOSPHATE + DNA(N).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF074328; AAD08667.1; -.
CC EMBL: AF084543; AAD2936.1; -.
DR

```

```

DR InterPro: IPR002064; -.
DR Pfam: PF00136; DNA_POL_B: 2.
DR PRINTS: PR00106; DNAPOLB.
DR PROSITE: PS00116; DNA_POLYMERASE_B: 1.
KW Transferrase: DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 1171 AA; 128590 MW; D2D64897FD5E70E8 CRC64;

```

```

Query Match          75.6%; Score 34; DB 1; Length 1171;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 GRLCYOD 7
   111111
Db 515 GRLCYOD 521

```

```

Search completed: June 13, 2001, 14:21:42
Job time: 802 sec

```


This page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:20:26 ; Search time 225.85 Seconds
(without alignments)
4.152 Million cell updates/sec

Title: PCT-US01-05825A-6

Perfect score: 45

Sequence: 1 GRLCVQDG 8

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_15:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organella:*
10: sp.phage:*
11: sp.plant:*
12: sp.podent:*
13: sp.unclassified:*
14: sp.vertebrate:*
14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	323	2	0918F5 vibrio mimi
2	45	100.0	399	2	09R3V6 vibrio chol
3	45	100.0	399	2	0917Q6 vibrio chol
4	38	84.4	549	11	09WU64 mus musculu
5	37	82.2	217	4	014964 homo sapien
6	36	80.0	870	3	014424 coccidioides
7	35	77.8	417	10	049465 mus musculu
8	35	77.8	611	2	069514 arabisidopsi
9	35	77.8	1677	11	070373 mycobacteri
10	34	75.6	296	5	070373 mus musculu
11	34	75.6	475	10	09XX28 caenorhabd
12	34	75.6	775	11	081325 arabidopsis
13	34	75.6	1048	14	069390 mus musculu
14	33	73.3	196	2	084353 pseudorabie
15	33	73.3	196	2	084353 chlamydia t
16	33	73.3	263	10	09PK45 chlamydia m
17	33	73.3	447	2	09ZQ11 arabidopsis
18	33	73.3	516	2	086673 streptomyc
19	33	73.3	516	2	084333 chlamydia t
					09PK65 chlamydia m

20	33	73.3	906	5	09V6Z2 drosophila
21	33	73.3	967	5	018965 caenorhabd
22	33	73.3	971	13	093248 oncorhynch
23	33	73.3	978	4	060313 homo sapien
24	33	73.3	1007	14	09IF17 porcine cyt
25	33	73.3	1007	14	09IF16 porcine cyt
26	33	73.3	1007	14	09IF15 porcine cyt
27	33	73.3	1007	14	09IF13 porcine cyt
28	33	73.3	1013	14	092827 human herpe
29	33	73.3	2108	2	053901 mycobacteri
30	33	73.3	2146	5	09VC97 xenopus lae
31	33	73.3	2327	13	09IBG7 drosophila
32	32	71.1	111	5	09VBN8 phytolecca
33	32	71.1	126	10	09S9F4 thiodaciliu
34	32	71.1	243	2	06168 drosophila
35	32	71.1	340	5	06VMS8 caenorhabd
36	32	71.1	440	5	076463 delinococcus
37	32	71.1	457	2	09RJ4 streptomyc
38	32	71.1	541	2	09RJ71 arabidopsis
39	32	71.1	809	10	09LP29 arabidopsis
40	32	71.1	830	10	09ZT06 arabidopsis
41	32	71.1	857	10	039392 brassica ol
42	32	71.1	870	3	013385 cochlidiobol
43	32	71.1	870	3	013391 cochlidiobol
44	32	71.1	984	9	080102 bacterioph
45	32	71.1	1461	11	09JLP3 mus musculu

ALIGNMENTS

RESULT 1

ID 09L8F5 PRELIMINARY; PRT; 323 AA.

AC 09L8F5;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE ZOR (FRAGMENT).

GN ZOR.

OS Vibrio mimicus.

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

OX NCBI_Taxid=674;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PT5;

RX MEDLINE=20143766; PubMed=10678967;

RA Boyd E.F., Moyer K.E., Shi L., Walder M.K.;

RT "Infectious CTXphi and the vibrio pathogenicity island prophage in vibrio mimicus: evidence for recent horizontal transfer between V.

RL mimicus and V. cholerae";

RU Infect. Immun. 68:1507-1513(2000).

DR EMBL; AF207857; AAF40142.1; -.

FT NON_TER 1

FT NON_TER 323

SO SEQUENCE 323 AA; 36306 MW; 01C12DAE9B873C3B CRC64;

Query Match 100.0%; Score 45; DB 2; Length 323;

Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRLCVQDG 8

Db 240 GRLCVQDG 247

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

RESULT 2

ID 09R3V6 PRELIMINARY; PRT; 399 AA.

AC 09R3V6;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

```

DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE ZONULAR OCCUDENS TOXIN (ZONA OCCUDENS TOXIN).
GN ZOT OR VC1458.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KNH002:
RA Shln H.J., Park Y.C., Kim Y.C.;
RT "Cloning and nucleotide sequence analysis of the virulence gene
RL MsaImumhag Holji 35:205-210(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-0139-TOR OGAWA:
RA Zhi-Yong H., Wei-Jie Z., Xiang-Fu W.;
RT "Cloning and Expression of zot Gene from Vibrio cholerae.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Hart D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Drigol I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL: AF175708; AAD51358.1; -
DR EMBL: AF123049; AAD26854.1; -
DR EMBL: AE004224; AAF94615.1; -
DR TIGR: VC1458; -
SO SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;

Query Match 100.0%; Score 45; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRICVODG 8
Db 291 GRICVODG 298

RESULT 3
ID 09L706 PRELIMINARY; PRT: 399 AA.
AC 09L706;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE ZOT.
GN ZOT.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=86015;
RA Kan B., Liu Y.Q., Qi G.M., Gao S.Y.;
RT "Vibrio cholerae nct-CTxphi whole genome, include rstr(Rstr),
RT rsta(Rsta), rslb(Rslb), cep(Cep), orfu(Orfu), ace(Ace) and zot(zot)
RT genes.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF220606; AAF29547.1; -
SO SEQUENCE 399 AA; 44930 MW; CF6A3DBCC9E23EE1 CRC64;

```

```

Query Match 100.0%; Score 45; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRICVODG 8
Db 291 GRICVODG 298

RESULT 4
ID 09W064 PRELIMINARY; PRT: 549 AA.
AC 09W064;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE GLYCEROL KINASE-LIKE PROTEIN 1.
GN GK-RS1 OR GKRS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1295SV; TISSUE=TESTES;
RA Pan Y., Decker W.K., Hug A.H.H.M., Craigen W.J.;
RT "Retroransposition of glycerol kinase-related genes from the X
RT chromosome to autosomes: functional and evolutionary aspects.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF117733; AAD24550.1; -
DR HSP: P08859; IGIC.
DR MGD: MGI:891990; Gk-rs1.
DR INTERPRO: IPR000577; -
DR INTERPRO: IPR002086; -
DR PFAM: PF00370; FSGY.1.
DR PROSITE: PS00070; ALDEHYDE-DEHYDR-CYS; UNKNOWN_1.
DR PROSITE: PS00445; FSGY_KINASES_2; 1.
DR PROSITE: PS00933; FSGY_KINASES_1; UNKNOWN_1.
KW kinase.
SO SEQUENCE 549 AA; 59871 MW; F807404B79F373F9 CRC64;

Query Match 84.4%; Score 38; DB 11; Length 549;
Best Local Similarity 75.0%; Pred. No. 9.7;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRICVODG 8
Db 266 GQLCLODG 273

RESULT 5
ID 014964 PRELIMINARY; PRT: 217 AA.
AC 014964;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE RAB-RELATED GTP-BINDING PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97237046; PubMed=9119394;
RA Stankovic T., Byrd P.J., Cooper P.R., McConville C.M., Munroe D.J.,
RA Riley J.H., Wats G.D., Ambrose H., McGuire G., Smith A.D.,
RA Sutcliffe A., Mills T., Taylor A.M.R.;
RT "Construction of a transcription map around the gene for ataxia
RT telangiectasia: identification of at least four novel genes.";
RL Genomics 40:267-276(1997).
DR EMBL: X99962; CAA68227.1; -

```

DR HSSP: P05713; 3RAB.
 DR INTERPRO: IPR001806; -.
 DR PFAM: PF00071; 188; 1.
 DR PRINTS: PRO00449; RASTRNSFRMG.
 FT NON_TER 1 1
 FT NON_TER 217 217
 SQ SEQUENCE 217 AA; 24869 MW; AA9A29731F42B8F5 CRC64;

Query Match 82.2%; Score 37; DB 4; Length 217;
 Best Local Similarity 62.5%; Pred. No. 6.4;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GRLCYODG 8
 ||:||||
 Db 182 GEICIQDG 189

RESULT 6
 014424 PRELIMINARY; PRT: 870 AA.

AC 014424;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE BETA-GLUCOSIDASE (EC 3.2.1.21).
 GN BGL.
 OS Coccidioides Immittis.
 OC Eukaryota; Fungi; Ascomycota; Onygenales; anamorphic Onygenales;
 CC Coccidioides.
 OX NCBI_TaxID=5501;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C735;
 RA Yu J.-J., Thomas P.W., Seshan K., Cole G.T.;
 RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U87805; AAB67972.1; -.
 DR INTERPRO: IPR001764; -.
 DR INTERPRO: IPR002772; -.
 DR PFAM: PF00933; Glyco_hydro_3; 1.
 DR PFAM: PF01915; Glyco_hydro_3_C; 1.
 DR PRINTS: PR00133; GLHYDRASE3.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
 KM Hydrolyase; Glycosidase.
 SO SEQUENCE 870 AA; 95467 MW; E0FDDEF20CD64E1F CRC64;

Query Match 80.0%; Score 36; DB 3; Length 870;
 Best Local Similarity 85.7%; Pred. No. 37;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLCYODG 8
 ||:||||
 Db 103 RLCIQDG 109

RESULT 7
 049465 PRELIMINARY; PRT: 417 AA.

AC 049465;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE HYPOHETICAL 47.7 KDA PROTEIN.
 GN F24J7.30 OR AT4G19470.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 CC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Vitale D., Liguori R., Argiriou A., De Simone V.,

RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Vitale D., Liguori R., Argiriou A., De Simone V., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AL021768; CAAL6924.1; -.
 DR EMBL: AL161550; CAB78949.1; -.
 DR MENDEL: 27536; Atatn13358; 27536.
 DR INTERPRO: IPR001611; -.
 DR PFAM: PF00560; LRR; 1.
 KM Hypothetical protein.
 SO SEQUENCE 417 AA; 47658 MW; 18A577B4C8D1A85 CRC64;

Query Match 77.8%; Score 35; DB 10; Length 417;
 Best Local Similarity 85.7%; Pred. No. 30;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRLCYOD 7
 ||:||||
 Db 373 GRLCYOE 379

RESULT 8
 069514 PRELIMINARY; PRT: 611 AA.

AC 069514;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE DNA POLYMERASE III SUBUNIT GAMMA AND TAU.
 GN DNAX.
 OS Mycobacterium leprae.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seeger K., Harris D.;
 RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Parthill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93188700; PubMed=8446027;
 RA Eigmeier K., Honore N., Woods S.A., Caudron B., Cole S.T.;
 RT "Use of an ordered cosmid library to deduce the genomic organization
 of Mycobacterium leprae."
 RL Mol. Microbiol. 7:197-206(1993).
 DR EMBL: AL023596; CAAL9155.1; -.
 DR INTERPRO: IPR000862; -.
 SO SEQUENCE 611 AA; 65535 MW; 39BAD5F9E63636F6 CRC64;

Query Match 77.8%; Score 35; DB 2; Length 611;
 Best Local Similarity 62.5%; Pred. No. 43;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GRLCYODG 8
 ||:||||
 Db 187 GRICAOEG 194

RESULT 9
 070373

ID 070373 PRELIMINARY; PRT; 1677 AA.
 AC 070373;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE XIN.
 GN XIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-CARDIAC MUSCLE;
 RA Wang D.-Z., Lin J.J.-C., Kitten G.T., Solursh M., Lin J.J.-C.;
 RL Front. Biosci. 1:0-0(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-CARDIAC MUSCLE;
 RA Wang D.-Z., Lin J.J.-C.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF051945; AAC06023.1;
 DR MGD; MGI:1333878; Xin.
 SO SEQUENCE 1677 AA; 182084 MW; A201CFC9A710C7FF CRC64;

Query Match 77.8%; Score 35; DB 11; Length 1677;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRICVODG 8
 DB 886 GSLCVDKG 893

RESULT 10
 09XX28 PRELIMINARY; PRT; 296 AA.
 AC 09XX28;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE Y39A1A.11 PROTEIN.
 GN Y39A1A.11.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-94150718; PubMed-7906398;
 RA Wall M.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-94150718; PubMed-7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).
 DR EMBL; AL031633; CAA21012.1;
 DR HSSP; P19992; 1HDC.

DR INTERPRO: IPR002198;
 DR INTERPRO: IPR002347;
 DR INTERPRO: IPR002424;
 DR PFAM: PF00106; adh_short.1.
 DR PFAM: PF00678; adh_short_C2.1.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PRINTS: PR00081; GDRDH..
 DR PRINTS: PR01167; INSADHFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
 SO SEQUENCE 296 AA; 31244 MW; B1518A53D4ABE025 CRC64;

Query Match 75.6%; Score 34; DB 5; Length 296;
 Best Local Similarity 75.0%; Pred. No. 34;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRICVODG 8
 DB 16 GALSVDHG 23

RESULT 11
 081325 PRELIMINARY; PRT; 475 AA.
 AC 081325;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE F6N15.2 PROTEIN.
 GN F6N15.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA WASHU;
 RT "The A. thaliana Genome Sequencing Project.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Ryan E., Edwards J., Pape K.;
 RT "The sequence of A. thaliana F6N15.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Waterston R.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF069299; AAC19301.1;
 SO SEQUENCE 475 AA; 53173 MW; 22A1951B3828A173 CRC64;

Query Match 75.6%; Score 34; DB 10; Length 475;
 Best Local Similarity 71.4%; Pred. No. 53;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 RLCVODG 8
 DB 208 RICEVDG 214

RESULT 12
 090X23 PRELIMINARY; PRT; 775 AA.
 AC 090X23;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE SEMAPHORIN M-SEMAK.


```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57/BL6K 6;
RA Miyazaki N., Furuyama T., Inagaki S.;
RT "A novel semaphorin, W-Semak which inhibits neural outgrowth from
RL sensory neurons."
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF034744; AAD01966.1; -
DR INTERPRO: IPR001627; -
DR INTERPRO: IPR003006; -
DR PFMAM: PF00047; 19; 1.
DR PFMAM: PF01403; Sema; 1.
SQ SEQUENCE 775 AA; 89542 MW; 221E766F404098D4 CRC64;

Query Match
Best Local Similarity 75.6%; Score 34; DB 11; Length 775;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRLCYOD 7
Db 267 GRLCYND 273

RESULT 13
O96390 PRELIMINARY; PRT; 1048 AA.
AC O96390;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE DNA POLYMERASE (EC 2.7.7.7) (FRAGMENT).
OS Pseudorabies virus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10345;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KAPLAN;
RX MEDLINE=95222727; PubMed=7707503;
RA Berthomme H., Monahan S.J., Parris D.S., Jacquemont B., Epstein A.L.;
RT "Cloning, sequencing, and functional characterization of the two
RT subunits of the pseudorabies virus DNA polymerase holoenzyme: evidence
RT for specificity of interaction."
RL J. Virol. 69:2811-2818(1995).
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE - N
CC PYROPHOSPHATE + DNA(N).
DR EMBL: L24487; AAA74383.1; -
DR INTERPRO: IPR002064; -
DR PFMAM: PF00136; DNA_POL_B; 1.
DR PRINTS: PR00106; DNAPOLB.
DR PROSITE: PS00116; DNAPOLYMERASE_B; 1.
KM DNA-directed DNA polymerase; DNA replication; DNA-binding.
FT NON_TER 1048 1048
SQ SEQUENCE 1048 AA; 115336 MW; B1EB70CFD389276D CRC64;

Query Match
Best Local Similarity 75.6%; Score 34; DB 14; Length 1048;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GRLCYOD 7
Db 475 GRLCYND 481

RESULT 14
O84353 PRELIMINARY; PRT; 196 AA.
ID O84353

```

```

AC O84353;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
DE MAF-TYPE PROTEIN.
GN MAF.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis."
RL Science 282:754-759(1998).
DR EMBL: AE001308; AAC67944.1; -
SQ SEQUENCE 196 AA; 21978 MW; E0B6CFED52F93073 CRC64;

Query Match
Best Local Similarity 73.3%; Score 33; DB 2; Length 196;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GRLCYODG 8
Db 150 GRLCYODG 157

RESULT 15
O9PK45 PRELIMINARY; PRT; 196 AA.
ID O9PK45;
AC O9PK45;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE MAF PROTEIN.
GN TC0628.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPN / N16G;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MOPN and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Read T.D., Brunham R., Shen C., Gill S.R., Heidelberg J.F., White O.,
RA Hickey E.K., Peterson J., Umayam L.A., Utterback T., Berry K.,
RA Bass S., Linher K., Weidman J., Khouri H., Craven B., Bowman C.,
RA Dodson R., Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G.,
RA Salzberg S.L., Eisen J., Fraser C.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE002331; AAF39457.1; -
DR TIGR: TC0628; -
SQ SEQUENCE 196 AA; 21730 MW; 9B86B059E52D78A5 CRC64;

```

```

Query Match
Best Local Similarity 73.3%; Score 33; DB 2; Length 196;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Wed Jun 13 15:01:16 2001

pct-us01-05825a-6.rspt

Page 6

Oy 1 GRUCVODG 8
| | | | |
Db 150 GGYCVODG 157

Search completed: June 13, 2001, 14:20:26
Job time: 726 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:16:32 ; Search time 118.55 Seconds
(without alignments)
1.296 Million cell updates/sec

Title: PCT-US01-05825A-6

Perfect score: 45

Sequence: 1 GRICVODG 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents-AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	400	2	US-08-624-601-8
2	36	80.0	745	1	US-08-453-472-5
3	36	80.0	745	1	US-08-038-948-9
4	36	80.0	745	1	US-08-453-952-5
5	36	80.0	745	2	US-08-484-993B-43
6	36	80.0	745	2	US-08-862-903-5
7	36	80.0	745	2	US-08-484-158B-43
8	36	80.0	745	2	US-08-484-596A-43
9	36	80.0	745	2	US-08-480-150A-43
10	36	80.0	745	3	US-08-458-731-43
11	36	80.0	745	3	US-08-149-223A-43
12	34	75.6	715	2	US-08-484-993B-10
13	34	75.6	715	2	US-08-484-158B-10
14	34	75.6	715	2	US-08-484-596A-10
15	34	75.6	715	2	US-08-480-150A-10
16	34	75.6	715	3	US-08-458-731-10
17	34	75.6	715	3	US-08-149-223A-10
18	34	75.6	716	2	US-08-484-993B-14
19	34	75.6	716	2	US-08-484-158B-14
20	34	75.6	716	2	US-08-484-596A-14
21	34	75.6	716	2	US-08-480-150A-14
22	34	75.6	716	3	US-08-458-731-14
23	34	75.6	716	3	US-08-149-223A-14
24	32	71.1	115	4	US-08-513-974B-38
25	32	71.1	132	4	US-08-513-974B-325
26	32	71.1	132	4	US-08-513-974B-326
27	32	71.1	328	3	US-08-459-046-2

28	32	71.1	328	4	US-08-513-974B-39	Sequence 39, Appl
29	32	71.1	328	4	US-08-513-974B-56	Sequence 56, Appl
30	32	71.1	328	4	US-08-513-974B-371	Sequence 371, App
31	32	71.1	328	4	US-08-513-974B-380	Sequence 380, App
32	31	68.9	190	1	US-08-106-981-2	Sequence 2, Appl
33	31	68.9	448	2	US-08-811-897A-18	Sequence 18, Appl
34	31	68.9	448	2	US-08-855-213-18	Sequence 18, Appl
35	31	68.9	467	2	US-08-811-897A-19	Sequence 19, Appl
36	31	68.9	467	2	US-08-855-213-19	Sequence 19, Appl
37	31	68.9	476	2	US-08-811-897A-20	Sequence 20, Appl
38	31	68.9	476	2	US-08-855-213-20	Sequence 20, Appl
39	31	68.9	495	2	US-08-811-897A-21	Sequence 21, Appl
40	31	68.9	495	2	US-08-855-213-21	Sequence 21, Appl
41	31	68.9	744	2	US-08-462-080B-2	Sequence 2, Appl
42	31	68.9	744	3	US-08-462-090-2	Sequence 2, Appl
43	31	68.9	744	3	US-08-463-461-2	Sequence 2, Appl
44	31	68.9	1242	2	US-08-680-326-33	Sequence 33, Appl
45	31	68.9	2629	2	US-08-751-189-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-624-601-8
Sequence 8, Application US/08624601
Patent No. 5862653
GENERAL INFORMATION:
APPLICANT: Kaper Dr., James B.
APPLICANT: Levine Dr., Myron M.
TITLE OF INVENTION: Vibrio cholerae O1 (CVD111) and non-O1
TITLE OF INVENTION: CVD112 and CVD112(RN) serogroup vaccine strains, methods
TITLE OF INVENTION: Of making same and products thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Spencer & Frank
STREET: 1100 New York Ave. N.W. Suite 300 East
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/624,601
FILING DATE: 08-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schenliet Dr., John W.
REGISTRATION NUMBER: 26,031
REFERENCE/DOCKET NUMBER: BAWCZ0019P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)414-4000
TELEFAX: (202)414-4040
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Vibrio cholerae
STRAIN: El Tor 7946
IMMEDIATE SOURCE:
CLONE: zot
US-08-624-601-8

Query Match 100.0%; Score 45; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRICVODG 8
DB 291 GRICVODG 298

RESULT 2

US-08-453-472-5
Sequence 5, Application US/08453472
Patent No. 5626846
GENERAL INFORMATION:
APPLICANT: DEAN, JURRIEN
TITLE OF INVENTION: CONTRACEPTIVE VACCINE
TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,472
FILING DATE: 30-May-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/038,948
FILING DATE: 26-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,462
FILING DATE: 20-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/364,379
FILING DATE: 12-JUN-1989
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4032 US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 745
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: ZP2
LOCATION:

IDENTIFICATION METHOD:
OTHER INFORMATION: human ZP2 protein
US-08-453-472-5

Query Match 80.0%; Score 36; DB 1; Length 745;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
DB 369 GELCTODG 376

RESULT 3

US-08-038-948-9
Sequence 9, Application US/08038948
Patent No. 5641487
GENERAL INFORMATION:
APPLICANT: DEAN, JURRIEN
TITLE OF INVENTION: CONTRACEPTIVE VACCINE BASED ON
TITLE OF INVENTION: ALLOIMMUNIZATION WITH ZONA PELLUCIDA POLYPEPTIDES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/038,948
FILING DATE: 26-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,462
FILING DATE: 20-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/364,379
FILING DATE: 12-JUN-1989
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: 99152/E-266-88/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-038-948-9

Query Match 80.0%; Score 36; DB 1; Length 745;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
DB 369 GELCTODG 376

RESULT 4

US-08-453-952-5
; Sequence 5, Application US/08453952
; Patent No. 5672488
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,952
; FILING DATE: 30-May-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,948
; FILING DATE: 26-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4032 US4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLER:
; FEATURE:
; NAME/KEY: ZP2
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: human ZP2 protein
; US-08-453-952-5

Query Match 80.0%; Score 36; DB 1; Length 745;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 GRLCVDG 8
| | | | |

Db 369 GELCTODG 376
RESULT 5
US-08-484-993B-43
; Sequence 43, Application US/08484993B
; Patent No. 5837497
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunoccontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,993B
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-993B-43

Query Match 80.0%; Score 36; DB 2; Length 745;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 GRLCVDG 8
| | | | |
Db 369 GELCTODG 376

RESULT 5
US-08-862-903-5
; Sequence 5, Application US/08862903
; Patent No. 5916768
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN

STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/862,903
FILING DATE: 30-May-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/038,948
FILING DATE: 26-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,462
FILING DATE: 20-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/364,379
FILING DATE: 12-JUN-1989
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AOTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4032 US4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 745
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: ZP2
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: human ZP2 protein
US-08-862-903-5

Query Match 80.0%; Score 36; DB 2; Length 745;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GRLCVQDG 8
| | | | |
Db 369 GELCTQDQ 376

RESULT 7
US-08-484-158B-43
Sequence 43, Application US/08484158B
Patent No. 5976545
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.

TITLE OF INVENTION: Pharmaceutical Compositions for
TITLE OF INVENTION: Immunoreception
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
CITY: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,158B
FILING DATE: 07-JUNE-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 09-NOV-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-92
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 32794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-158B-43

Query Match 80.0%; Score 36; DB 2; Length 745;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GRLCVQDG 8
| | | | |
Db 369 GELCTQDQ 376

RESULT 8
US-08-484-596A-43
Sequence 43, Application US/08484596A
Patent No. 5981228
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunoreception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,596A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 11-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-596A-43

Query Match 80.0%; Score 36; DB 2; Length 745;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GRICVQDG 8
| | | | |
DB 369 GELCTQDG 376

RESULT 9
US-08-480-150A-43
Sequence 43, Application US/08480150A
Patent No. 5989550
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,150A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/149,223
FILING DATE: 09-NOV-1993
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341

FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-150A-43

Query Match 80.0%; Score 36; DB 2; Length 745;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVQDG 8
| | | | |
DB 369 GELCTQDG 376

RESULT 10
US-08-458-731-43
Sequence 43, Application US/08458731
Patent No. 6001599
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,731
FILING DATE: 09-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-458-731-43

Query Match 80.0%; Score 36; DB 3; Length 745;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
1 1 1 1 1
DB 369 GELCTODG 376

RESULT 11
US-08-149-223A-43
; Sequence 43, Application US/08149223A
; Patent No. 6027727
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,223A
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-149-223A-43

Query Match 80.0%; Score 36; DB 3; Length 745;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
1 1 1 1 1
DB 369 GELCTODG 376

RESULT 12
US-08-484-993B-10
; Sequence 10, Application US/08484993B

; Patent No. 5837497
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,993B
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 715 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-993B-10

Query Match 75.6%; Score 34; DB 2; Length 715;
Best Local Similarity 75.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
1 1 1 1 1
DB 366 GDLCTODG 373

RESULT 13
US-08-484-158B-10
; Sequence 10, Application US/08484158B
; Patent No. 5976545
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Pharmaceutical Compositions for
; IMMUNOCONTRACEPTION
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America

ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,158B
FILING DATE: 07-JUNE-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 09-NOV-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-92
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 32794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 715 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-158B-10

Query Match 75.6%; Score 34; DB 2; Length 715;
Best Local Similarity 75.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
DB 366 GDLCTODG 373

RESULT 14
US-08-484-596A-10
Sequence 10, Application US/08484596A
Patent No. 5981228
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,596A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223

FILING DATE: 11-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 715 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-596A-10

Query Match 75.6%; Score 34; DB 2; Length 715;
Best Local Similarity 75.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVODG 8
DB 366 GDLCTODG 373

RESULT 15
US-08-480-150A-10
Sequence 10, Application US/08480150A
Patent No. 5989550
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,150A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/149,223
FILING DATE: 09-NOV-1993
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 715 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-150A-10

Query Match 75.68; Score 34; DB 2; Length 715;
Best Local Similarly 75.08; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRICVDPG 8
Db 366 GDICTDPG 373

Search completed: June 13, 2001, 14:16:33
Job time: 494 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:14:33 ; Search time 229.28 Seconds
(without alignments)
1.995 Million cell updates/sec

Title: PCT-US01-05825A-7

Perfect score: 41

Sequence: 1 GRLLVQPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A:Geneseq_0401:*

1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT:*

2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT:*

3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT:*

4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT:*

5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT:*

6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT:*

7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT:*

8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT:*

9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT:*

10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT:*

11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT:*

12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT:*

13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT:*

14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT:*

15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT:*

16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT:*

17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT:*

18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT:*

19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT:*

20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT:*

21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT:*

22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	8	21	Y79111
2	38	92.7	8	21	Y79107
3	36	87.8	8	21	Y79109
4	36	87.8	118	18	W19878
5	36	87.8	314	21	G20127
6	36	87.8	321	21	G20126
7	36	87.8	461	21	G20125
8	35	85.4	314	21	G48786
9	35	85.4	321	21	G48785
10	35	85.4	461	21	G48784
11	35	85.4	599	20	Y37055

12	35	85.4	1122	20	Y26579	Murine telomerase
13	34	82.9	8	21	Y79123	Peptide antagonist
14	34	82.9	538	20	Y34750	Amino acid sequenc
15	33	80.5	8	21	Y79105	Peptide antagonist
16	33	80.5	8	21	Y79112	Peptide antagonist
17	33	80.5	8	21	Y79115	Peptide antagonist
18	33	80.5	246	14	R40924	Protein able to bi
19	33	80.5	525	17	R96103	Hepatitis E virus
20	33	80.5	525	17	R96104	Hepatitis E virus
21	33	80.5	540	17	R96101	Hepatitis E virus
22	33	80.5	540	17	R96102	Hepatitis E virus
23	33	80.5	547	13	R26189	Epidemic NANBH vir
24	33	80.5	549	17	R96091	Hepatitis E virus
25	33	80.5	549	17	R96092	Hepatitis E virus
26	33	80.5	549	19	R96090	Hepatitis E virus
27	33	80.5	659	14	R38787	Hepatitis E virus
28	33	80.5	659	14	R39308	HEV ORF2 protein.
29	33	80.5	659	18	W35827	Mexico strain HEV
30	33	80.5	659	20	W93387	Hepatitis E virus
31	33	80.5	660	12	R14619	Human HEV ORF 2 pr
32	33	80.5	660	14	R38785	Protein encoded by
33	33	80.5	660	14	R39306	HEV ORF2 protein.
34	33	80.5	660	15	R51265	Burma strain HEV. O
35	33	80.5	660	16	R70323	HEV strain protein
36	33	80.5	660	17	R91814	Hepatitis E virus
37	33	80.5	660	17	R96089	Hepatitis E virus
38	33	80.5	660	17	R96090	Hepatitis E virus
39	33	80.5	660	18	W35826	Hepatitis E virus
40	33	80.5	660	19	W81520	Hepatitis E virus
41	33	80.5	660	19	W80197	Protein encoded by
42	33	80.5	660	19	W76369	Hepatitis E virus
43	33	80.5	660	19	W71210	Protein encoded by
44	33	80.5	660	20	Y31385	HEV-US2 ORF2 prote
45	33	80.5	660	20	Y31382	HEV-US1 ORF2 prote

ALIGNMENTS

RESULT	ID	Y79111 standard; Peptide; 8 AA.	ALIGNMENTS
1	Y79111		
XX	AC	Y79111:	
XX	XX	05-JUN-2000 (first entry)	
XX	XX	Peptide antagonist of zonulin.	
DE	XX		
XX	KW	zonulin: antagonist; zonula occludens toxin receptor;	
KW	KW	blood-brain barrier; antinflammatory; cerebroprotective;	
KW	KW	neuroprotective; dermatological; antitumor; antiviral;	
KW	KW	antibacterial; cytostatic; anti-HIV; vulnereary; antiallergic;	
KW	KW	hypotensive; immunosuppressive; antiparasitic; vasotropic;	
KW	KW	gastrointestinal inflammation; therapy.	
XX	OS	Synthetic.	
XX	PN	WO200007609-A1.	
XX	XX		
PD	XX	17-FEB-2000.	
XX	PF	28-JUL-1999; 99WO-US16683.	
XX	XX		
XX	PR	03-AUG-1998; 98US-0127815.	
XX	PA	(UWMA-) UNIV MARYLAND BALTIMORE.	
XX	PI	Fasano A;	
XX	DR	WPI; 2000-205565/18.	
XX	PT	New peptide antagonist of zonulin useful as antinflammatory agent for	

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 XX
 PS Claim 1; Page 43; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulin, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC celiac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

CC Sequence 8 AA:

Query Match 100.0%; Score 41; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRLLVOPG 8
 ||:|||||
 Db 1 grllvpg 8

RESULT 2
 Y79107
 ID Y79107 standard; Peptide; 8 AA.

XX Y79107;

DT 05-JUN-2000 (first entry)

XX Peptide antagonist of zonulin.

XX Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antilucer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.

XX Synthetic.

PN WO200007609-A1.

XX 17-FEB-2000.

XX 28-JUL-1999; 99WO-US16683.

XX 03-AUG-1998; 98US-0127815.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

PI Fasano A;
 XX WPI; 2000-205565/18.

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 XX
 PS Claim 1; Page 41; 69pp; English.

CC This present sequence is that of a peptide antagonist of zonulin
 CC (2), one of 25 such peptides (see Y79105-29) of the invention,
 CC which bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulin, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The peptide antagonists are used as an
 CC antiinflammatory agents in the treatment of gastrointestinal
 CC inflammation, where they bind to the ZOT receptor in the intestine
 CC and yet does not physiologically modulate the opening of TJ in the
 CC intestine. Gastrointestinal inflammation conditions give rise to
 CC increased intestinal permeability and the peptide is useful for
 CC treating intestinal conditions that cause protein losing enteropathy
 CC caused by infection, e.g. Clostridium difficile infection,
 CC enterocolitis, shigellosis, viral gastroenteritis, parasite
 CC infestation, bacterial overgrowth, whipple's disease, diseases with
 CC mucosal erosion or ulcerations, e.g. gastritis, gastric cancer,
 CC collagenous colitis, inflammatory bowel disease, diseases marked by
 CC lymphatic obstruction, e.g. congenital intestinal lymphangiectasia,
 CC sarcoidosis lymphoma, mesenteric tuberculosis, and after surgical
 CC correction of congenital heart disease with Fontan's operation,
 CC mucosal diseases without ulceration, e.g. Menetrier's disease,
 CC celiac disease, eosinophilic gastroenteritis, and immune diseases,
 CC e.g. systemic lupus erythematosus or food allergies, primarily to
 CC milk.

CC Sequence 8 AA:

Query Match 92.7%; Score 38; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 3.2e+05;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRLLVOPG 8
 ||:|||||
 Db 1 grllvpg 8

RESULT 3
 Y79109
 ID Y79109 standard; Peptide; 8 AA.

XX Y79109;

DT 05-JUN-2000 (first entry)

XX Peptide antagonist of zonulin.

XX Zonulin; antagonist; zonula occludens toxin receptor;
 KW blood-brain barrier; antiinflammatory; cerebroprotective;
 KW neuroprotective; dermatological; antilucer; antiviral;
 KW antibacterial; cytostatic; anti-HIV; vulnery; antiallergic;
 KW hypotensive; immunosuppressive; antiparasitic; vasotropic;
 KW gastrointestinal inflammation; therapy.

XX Synthetic.

PN WO200007609-A1.

XX 17-FEB-2000.

XX 28-JUL-1999; 99WO-US16683.